



Kent Academic Repository

Lan, Hai, Luo, Yong (Eddie), Lowe, Ben, Gong, Yongzhi and Tang, Xiaofei (2026) *Tackling tricky complaints: the impact of AI agents and intention hiding strategies on user responses*. Internet Research . ISSN 1066-2243.

Downloaded from

<https://kar.kent.ac.uk/113109/> The University of Kent's Academic Repository KAR

The version of record is available from

<https://doi.org/10.1108/INTR-11-2024-1831>

This document version

Author's Accepted Manuscript

DOI for this version

Licence for this version

CC BY-NC (Attribution-NonCommercial)

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in **Title of Journal**, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

Tackling tricky complaints: the impact of AI agents and intention hiding strategies on user responses

Hai Lan¹

Southwestern University of Finance and Economics

Yong (Eddie) Luo

University of Kent

Ben Lowe

University of Kent

Yongzhi Gong

Southwestern University of Finance and Economics

Xiaofei Tang

Southwestern University of Finance and Economics

This is an Accepted Manuscript of an article published in *Internet Research*

Please cite: Lan, Hai, Yong (Eddie) Luo, Ben Lowe, Yongzhi Gong and Xiaofei Tang (2026), "Tackling tricky complaints: the impact of AI agents and intention hiding strategies on user responses," Forthcoming *Internet Research*, (Published online 7th January 2026). DOI: **TBD**

Available online at: <http://dx.doi.org/???>

¹ Corresponding author. All correspondence should be addressed to lanhai@smail.swufe.edu.cn

Tackling tricky complaints: the impact of AI agents and intention hiding strategies on user responses

Abstract

Purpose – This research serves a twofold purpose: first, to identify and categorize two common intention hiding strategies used by frontline employees when handling tricky user complaints, specifically evasive hiding and rationalized hiding; and second, to systematically examine the interactive effects of agent type (AI vs. human) and these strategies on users' willingness to forgive.

Design/methodology/approach – Three experiments (N = 820) were conducted to investigate how agent type (AI vs. human) interacts with different intention hiding strategies to influence users' willingness to forgive. The experiments also tested the mediating effects of perceived negative motives and perceived sincerity, exploring how AI capabilities (mechanical vs. thinking) shape user reactions.

Findings – Users exhibit a higher willingness to forgive AI agents than human agents when an evasive hiding strategy is used; conversely, human agents elicit more favorable responses when employing a rationalized hiding strategy. These effects are mediated by perceived negative motives and perceived sincerity. Furthermore, mechanical AI agents are more effective when using evasive hiding strategies, whereas thinking AI agents perform better with rationalized hiding strategies.

Originality/value – This research extends service recovery theory by introducing evasive and rationalized hiding as intention hiding strategies and by demonstrating that user responses vary according to the alignment between agent type and hiding strategy type. The findings also enrich research on mind perception and AI interaction by uncovering the underlying psychological processes and highlighting the influence of AI capability design on users' interpretations.

Keywords tricky complaints, intention hiding strategies, AI agent, willingness to forgive

Introduction

In service settings, frontline employees sometimes encounter tricky situations where they know the company has no intention of addressing user complaints or resolving the issue in the short term, yet they are still required to respond to customer inquiries or complaints (Muzatko and Bansal, 2023). Such “must-respond but cannot genuinely resolve” complaints and queries can be characterized as tricky user complaints. To maintain relationships and reduce conflict without issuing a direct refusal, companies often instruct employees to deliberately conceal the organization’s true intentions (Anand *et al.*, 2022; Gong *et al.*, 2025) and to employ response phrasing that appears to address the issue while effectively avoiding substantive resolution.

For example, consider *Dongqiudi*, a Chinese mobile application that provides various football-related information services. In response to users’ questions about why the application had not yet introduced a night mode despite repeated requests, customer service representatives, who were aware that the company had no intention of addressing the issue, typically replied with, “Thank you for your suggestion. We will pass it on to the relevant team and consider it in future versions.” Similarly, in response to another user’s complaint about the excessive presence of content irrelevant to football in personalized recommendations, the customer service representative explained, “To attract a broader user base, we have recently expanded our content categories.” Although these responses appear polite on the surface, they convey little substantive information. Instead, they strategically employ linguistic ambiguity to obscure the company’s inaction.

To frame the inquiry of the current article, the authors note that these two types of responses represent the linguistic strategies used to conceal an organization’s true intentions when dealing with tricky user complaints, namely evasive hiding and rationalized hiding (Connelly *et al.*, 2012; Gong *et al.*, 2025). The former relies on vague commitments (e.g., “We will consider it in the future”) to create an illusion of

positive responsiveness, while the latter appeals to external constraints or third-party responsibilities (e.g., “In consideration of broader user needs”), implying that the issue cannot be resolved in the short term. The focus of this research is not on evaluating which strategy is better or worse. In many cases, the wording used to address user complaints is pre-specified by business managers, and frontline agents simply follow established protocols.

Considering the rapid advancement of AI technologies, another aspect related to the current research framework is the use of AI agents to improve efficiency and reduce costs (Bergner et al., 2023; Xie et al., 2025a). However, user evaluations of AI agents and human agents differ significantly (Jin et al., 2025; Lan et al., 2025). Although prior research has examined the use of positive affective strategies, such as humor and empathy, in AI-based service recovery contexts (Han *et al.*, 2023; Liu *et al.*, 2023; Lv *et al.*, 2022), the findings remain inconclusive when AI agents are directly compared with human agents. Some studies suggest that AI may elicit stronger negative reactions (Longoni *et al.*, 2023), whereas others indicate that AI agents are more likely to be granted leniency (Srinivasan and Sarial-Abi, 2021). Existing research has not yet incorporated hiding strategies that are widely used in practice, such as evasive hiding and rationalized hiding, nor has it explored users’ cognitive differences when AI agents employ these strategies. To address this gap and advance theoretical understanding, the present research investigates how agent type (AI vs. human) and intention hiding strategy (evasive vs. rationalized) jointly shape users’ willingness to forgive the company in tricky feedback situations.

Grounded in mind perception theory (Gray *et al.*, 2007), this research explores the impact of agent type and intention hiding strategy on users’ willingness to forgive through two psychological mechanisms: perceived negative motive and perceived sincerity. While perceived negative motive pertains to the inference of an agent’s underlying intent, perceived sincerity relates to the perceived genuineness of emotional displays (Antonetti and Baghi, 2023; Béal and Grégoire, 2022). Given that

systematic variations in mind perception between AI and humans can spill over into firm-level evaluations (Garvey *et al.*, 2023; Lan *et al.*, 2025), we argue that different agent-strategy combinations elicit divergent mind perceptions that ultimately determine the willingness to forgive. Finally, drawing on the AI capability framework (Huang and Rust, 2021), we examine users' willingness to forgive AI agents under different intention hiding strategies by assigning AI agents distinct capability labels (mechanical vs. thinking), thereby deepening understanding of the application of AI in service recovery contexts.

Overall, this research is poised to make several theoretical contributions. First, it introduces evasive hiding and rationalized hiding, as two distinct intention hiding strategies, into the service recovery literature, extending existing frameworks beyond compensatory and affective responses. Second, it advances understanding of AI in service recovery by theorizing how agent type interacts with intention hiding strategies to shape user forgiveness, thereby transcending the focus on strategies alone to incorporate the role of the agents adopting the strategies. Finally, by integrating mind perception theory and AI capability distinctions, the research offers a process-oriented perspective on how cognitive and affective inferences jointly drive user responses in AI-mediated service interactions.

Literature review

AI agents handling tricky user complaints

AI agents are increasingly pivotal in managing service recovery, although inherent limitations in emotional and contextual depth persist when addressing tricky user complaints (Longoni *et al.*, 2023; Liu *et al.*, 2023). While anthropomorphism may intensify user dissatisfaction in high-stakes failure scenarios (Choi *et al.*, 2020; Fan *et al.*, 2020; Srinivasan and Sarial-Abi, 2021; Crollic *et al.*, 2022), research also shows that recovery success can be achieved through “cute” cues, multimodal empathy, and strategic linguistic tactics (Lv *et al.*, 2021; Lv *et al.*, 2022; Liu *et al.*,

2023). For example, self-deprecating humor can significantly increase the willingness to forgive (Xie *et al.*, 2025b). However, the efficacy of these expressions depends on both their intensity and their alignment with the task's complexity and the user's emotions (Chen *et al.*, 2025; Liu and Xu, 2023).

Existing literature often treats AI and human agents as functionally equivalent, overlooking the distinct psychological attributions users form based on an agent's perceived identity (Xie *et al.*, 2025b; Garvey *et al.*, 2023; Lan *et al.*, 2025). Furthermore, while previous studies have focused on positive emotional responses, they largely ignore how agents manage the most challenging feedback through intention hiding strategies, such as evasive or rationalized hiding (Gong *et al.*, 2025). These strategies are particularly relevant for handling tricky complaints, where organizations face resource or policy limitations, and reflect actual communication practices more accurately than purely emotional responses. Accordingly, studies on intention hiding strategies are essential to understanding how AI agents navigate complex, sensitive user feedback in real-world service environments.

Intention hiding strategies: evasive hiding vs. rationalized hiding

In service practice, firms and service agents are not always willing or able to provide truthful information or undertake substantive recovery efforts (Béal and Grégoire, 2022). Drawing on the organizational behavior literature, Connelly *et al.* (2012) conceptualized such behavior as “knowledge hiding,” defined as the deliberate withholding of requested information. Two intention hiding strategies (evasive hiding and rationalized hiding) from this framework were particularly relevant in service contexts where employees seek to preserve relationships, avoid conflict, or deflect responsibility (Chen *et al.*, 2024a; Michalová *et al.*, 2024). Although initially developed for internal organizational settings, recent research shows that these strategies are also prevalent in firm-customer interactions, especially when frontline employees are aware that the organization has no intention of addressing customer requests in the short term (Gong *et al.*, 2025). In such situations, service agents often

rely on specific linguistic responses to obscure the organization’s true intentions rather than explicitly rejecting the request.

Adapting the definitions of Connelly *et al.* (2012) to service interaction contexts, this research defines evasive hiding as the use of vague or misleading promises to create a sense of responsiveness despite no genuine intention to act, and rationalized hiding as the use of justificatory explanations that invoke external constraints or third-party considerations to legitimize inaction. While both strategies involve intention concealment, they differ in linguistic structure, transparency, and expectation management. Evasive hiding relies on fictitious future commitments and high ambiguity, whereas rationalized hiding offers partial transparency by shifting the focus from the company’s willingness to act to the feasibility of the action. Unlike common service recovery strategies such as apologies, compensation, empathy, humor, or low-competence apologies, intention hiding strategies do not aim to assume responsibility, regulate emotions, or incur economic or symbolic costs. Instead, they create informational asymmetries that shape customers’ inferences about organizational motives and future behavior (Antonetti and Baghi, 2023; Béal and Grégoire, 2022; Javornik *et al.*, 2020; Liu and Xu, 2023; Lv *et al.*, 2022; Ohtsubo and Watanabe, 2009; Wang *et al.*, 2023). The distinctions between evasive hiding, rationalized hiding, and other service recovery strategies are shown in Table 1.

[Insert Table 1]

Theoretical underpinning and hypotheses development

Agent types, hiding strategies, and user willingness to forgive

Users perceive AI and human agents differently based on the nature of the task and emotional intensity (Lan *et al.*, 2025; Longoni and Cian, 2022), and these differences shape how users attribute motives in intention hiding contexts. An evasive hiding strategy employs vague language to deflect accountability. In interactions with human agents, such responses may be perceived as deliberate avoidance of genuine

communication, thereby increasing users' doubt and reducing their willingness to forgive. By contrast, when the same evasive language is employed by an AI agent, users are likely to interpret the response as a system-driven or templated output generated by automated protocols (Srinivasan and Sarial-Abi, 2021; Xie *et al.*, 2025a). This interpretation mitigates frustration and disappointment, thereby leading to forgiveness in AI-delivered interactions.

Conversely, a rationalized hiding strategy invokes external constraints or institutional rules to justify the absence of remedial action (Connelly *et al.*, 2010; Gong *et al.*, 2025). Although it provides a moderate level of information transparency, its effectiveness relies on the perceived credibility and conversational responsiveness of the agent. Human agents are typically viewed as more contextually appropriate and credible when delivering such explanations, thereby facilitating user acceptance and increasing the willingness to forgive. In comparison, similar rationalized explanations delivered by AI agents are more likely to be interpreted by users as formulaic excuses. As a result, the weaker perceived effectiveness of these explanations diminishes users' willingness to forgive. Accordingly, it is hypothesized that:

H1a: Users will exhibit a higher willingness to forgive when an evasive hiding strategy is employed by an AI agent (vs. human).

H1b: Users will exhibit a higher willingness to forgive when a rationalized hiding strategy is employed by a human agent (vs. AI).

The mediating role of perceived negative motives and perceived sincerity

This research adopts mind perception theory (Gray *et al.*, 2007) to explore the psychological mechanisms underlying the differential effects of the hiding strategies employed by AI and human agents. This theory posits that individuals evaluate others, including non-human agents, along two key dimensions: agency, referring to the capacity for intention and action, and experience, referring to the ability to perceive and express emotions.

In the evasive hiding condition, where the agent avoids providing a clear explanation or commitment (e.g., saying “we will consider this in the future”), users may interpret the vague language as an attempt to maintain responsiveness while knowing that the problem will not be addressed. This leads users to infer negative motives on the part of the company (Béal and Grégoire, 2022). However, users do not always directly associate the content of the message with the company’s intent; rather, they consider the messenger’s characteristics when interpreting the underlying psychological states (Candrian and Scherer, 2022). As AI agents are typically perceived as lacking agency, that is, they do not possess subjective intentions or autonomous decision-making (Gray *et al.*, 2007; Holthöwer and van Doorn, 2023; Kim *et al.*, 2022), users are less likely to attribute negative motives to AI itself, even when the conveyed message signals the company’s self-interest. This results in reduced suspicion of the strategy and a higher willingness to forgive.

By contrast, in the context of rationalized hiding, service agents explain their inability to address users’ requests by invoking external constraints. We argue that this strategy is intrinsically linked to perceived sincerity. First, rationalized hiding essentially involves the disclosure of limited capability. Firms typically strive to show superior problem-solving competence to the public (Güntürkün *et al.*, 2020). However, acknowledging the inability to respond due to policy or resource constraints conflicts with the goal of showing strong competence. Such self-weakened behavior imposes a substantial cost on the firm, as it builds the legitimacy of the explanation by revealing its own limitations (Chaudhry and Loewenstein, 2019). Prior research has shown that when communicators are willing to convey truthful information at the expense of their own interests or image, users infer higher levels of benevolence and honesty (Ohtsubo and Watanabe, 2009). Accordingly, explanations grounded in this strategy enable users to believe that the reasons provided by service agents are not evasive but instead reflect a candid disclosure of genuine constraints, thereby enhancing perceived sincerity.

However, this underlying process is more salient in the context of human agents. Human agents are perceived to possess higher mental capacity and are therefore assumed to be aware of the risks associated with admitting inability. As a result, their explanations are more likely to be interpreted as sincere expressions of unavoidable constraints. In contrast, rationalized explanations offered by AI agents are likely to be perceived as preprogrammed logical routines, and statements of inability are interpreted merely as functional boundaries of the system rather than as costly self-disclosures. The absence of perceived significant costs for AI agents limits the capacity of rationalized hiding to enhance perceived sincerity in AI contexts. Accordingly, it is hypothesized that:

H2a: Perceived negative motives mediate the positive effect of AI agents' use of an evasive hiding strategy on users' willingness to forgive.

H2b: Perceived sincerity mediates the positive effect of human agents' use of a rationalized hiding strategy on users' willingness to forgive.

The influences of AI capability on users' responses to intention hiding strategies

This research further explores whether variations in AI capability design influence user responses to hiding strategies. Existing studies indicate that user reactions to AI are influenced by factors such as perception of the AI's intelligence and personality traits (Cheng *et al.*, 2024; Kim *et al.*, 2022; Pitardi *et al.*, 2024; Schepers *et al.*, 2022). According to the AI capability framework (Huang and Rust, 2021), AI can be classified into two types: mechanical AI and thinking AI. Mechanical AI primarily handles basic, repetitive tasks and is generally perceived by users as a low-capability tool with decision-making autonomy. In contrast, thinking AI is highly intelligent, capable of learning from data and adapting through algorithmic models, and is particularly skilled at managing complex and personalized tasks. Consequently, thinking AI tends to be perceived as more human-like. This classification aligns with business practices in AI design and reflects users' psychological expectations of AI capabilities.

Based on this classification, we propose that AI capability interacts with intention hiding strategies in affecting users' willingness to forgive. Specifically, evasive hiding employs vague language (e.g., "we will consider this in the future"), implying a delay in problem-solving rather than immediate action. For thinking AI, users may believe that the AI possesses sufficient capability to solve the problem but intentionally avoids responsibility, leading to stronger dissatisfaction and disappointment (Crollic *et al.*, 2022; Peter *et al.*, 2025). For mechanical AI, the same evasive hiding strategy is more likely to be attributed to the AI's limited ability rather than to deliberate avoidance, resulting in greater tolerance (Cheng, 2023).

In the context of rationalized hiding, the agent provides explanations of external constraints such as platform policies or regulatory requirements, communicating the intention to explain rather than offering a concrete solution. For thinking AI, the explanation is more likely to be perceived as reasonable, as such agents are associated with rationality, logic, and data-driven problem-solving capability. This perception leads users to believe that the AI is engaging in sincere and reasonable communication rather than evasively avoiding responsibility (Xie *et al.*, 2025a). By contrast, for mechanical AI, the same strategy may prompt users to question its legitimacy due to the agent's limited communicative capability and weaker ability to convey empathy. Accordingly, it is hypothesized that:

H3a: Users will exhibit a higher willingness to forgive when an evasive hiding strategy is employed by a mechanical AI (vs. thinking AI).

H3b: Users will exhibit a higher willingness to forgive when a rationalized hiding strategy is employed by a thinking AI (vs. mechanical AI).

The research model is shown in Figure 1.

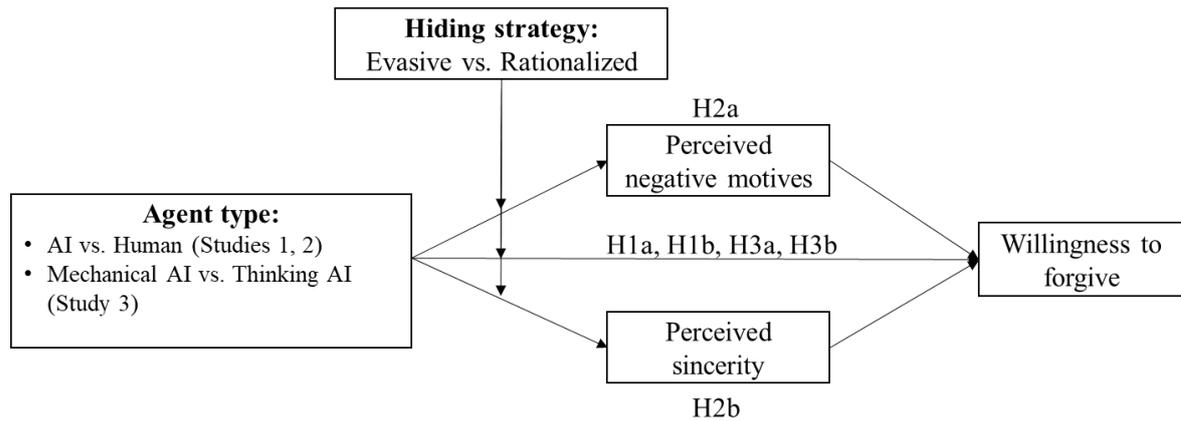


Figure 1. The research model
Source(s): Authors' own work

Study 1

Study 1 aimed to examine the effects of agent type and hiding strategy on users' willingness to forgive in a service interaction context. To enhance realism, the study adopted a common tricky complaints scenario, namely, the delayed update of a mobile application.

Method

Participants and design.

Study 1 employed a 2 (agent type: AI vs. human) × 2 (hiding strategy: evasive vs. rationalized) between-subjects design. A total of 306 participants (151 women, $M_{age} = 31.60$, $SD = 8.85$) were recruited from Credamo, a widely used research and data collection platform.

Procedure.

Participants were randomly assigned to one of the four experimental conditions. They were asked to imagine using a popular social app at night, which they found the screen's bright light was harsh and caused eye fatigue. They were then prompted to provide the following feedback regarding this issue: "Why doesn't this app have a night mode yet? Can't you learn from other apps?" In the AI (human) condition,

participants were informed that their feedback would be processed by an AI (human) agent. Agent type was manipulated through visual appearance cues following Garvey *et al.* (2023) (see Appendix A for details).

Next, participants read a response from the agent. In the evasive hiding conditions, the agent responded, “Hello, thank you for your feedback. We will evaluate this issue and strive to improve it in future versions.” In the rationalized hiding conditions, the agent responded, “Hello, thank you for your feedback. Considering compatibility issues with the technical architecture and prioritization of development resources, implementing night mode does present significant challenges in the short term.” After reading the materials, participants were asked to indicate whether the agent was a human or an AI agent. They also evaluated the agent’s hiding strategies using items adapted from Connelly *et al.* (2012). Subsequently, participants rated their willingness to forgive (Wei and Ran, 2019). Finally, participants responded to demographic questions. All measurement items were provided in Appendix B, and the reliability and validity assessment results were reported in Appendix C.

Results

Manipulation check.

The results indicated that participants in the evasive hiding conditions scored significantly higher on the evasive hiding measure compared to those in the rationalized hiding conditions ($M_{\text{Evasive}} = 4.65$, $SD = 1.58$; $M_{\text{Rationalized}} = 3.35$, $SD = 1.32$; $F(1, 304) = 59.70$, $p < 0.001$), while participants in the rationalized hiding conditions scored significantly higher on the rationalized hiding measure compared to those in the evasive hiding conditions ($M_{\text{Rationalized}} = 5.94$, $SD = 0.75$; $M_{\text{Evasive}} = 4.02$, $SD = 1.31$; $F(1, 304) = 243.10$, $p < 0.001$).

Willingness to forgive.

The results of a 2 (agent type: AI vs. human) \times 2 (hiding strategy: evasive vs. rationalized) ANOVA revealed a significant interaction effect ($F(1, 302) = 14.34$, $p <$

0.001) and a significant main effect of hiding strategy ($F(1, 302) = 21.29, p < 0.001$). The main effect of agent type was not significant ($F(1, 302) = 0.65, p > 0.05$). A simple effect analysis showed that participants demonstrated a higher willingness to forgive when the evasive hiding strategy was employed by AI agents than human agents ($M_{AI} = 5.32, SD = 1.14; M_{Human} = 4.88, SD = 1.27; F(1, 302) = 4.49, p < 0.05$). Conversely, participants showed a higher willingness to forgive when the rationalized hiding strategy was employed by human agents compared to AI agents ($M_{Human} = 4.76, SD = 1.32; M_{AI} = 4.08, SD = 1.37; F(1, 302) = 10.43, p < 0.01$) (see Figure 2). Thus, the results supported H1a and H1b.

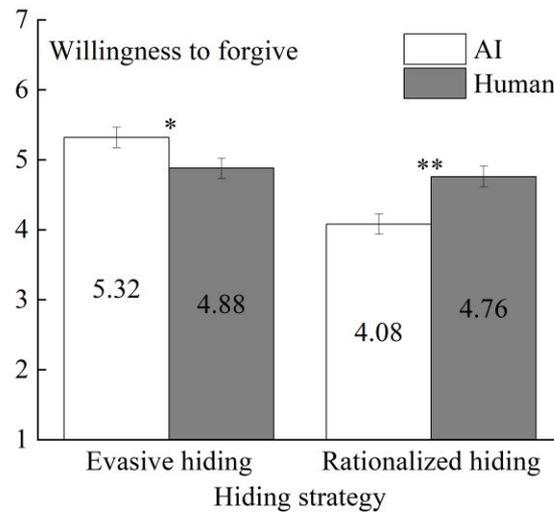


Figure 2. Study 1: Willingness to forgive as a function of agent type and hiding strategy

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; Error bars = ± 1 SEs.

Source(s): Authors' own work

Study 1 confirmed the interaction effect of agent type and hiding strategy on users' willingness to forgive when firms handle tricky complaints. Specifically, in the evasive hiding condition, users exhibited a higher willingness to forgive AI agents, whereas in the rationalized hiding condition, users showed a higher willingness to forgive human agents. Study 2 was designed to further investigate the underlying mechanisms of this effect.

Study 2

The objective of Study 2 was to test the proposed underlying processes and to examine H2. Simultaneously, we tested alternative explanations, including clarity, credibility, benevolent inference, and selfish inference, to rule out their potential effects (Antonetti and Baghi, 2023; Garvey *et al.*, 2023). To enhance the robustness of the findings, we utilized a different tricky complaint scenario involving software bugs.

Method

Participants and design.

Study 2 employed a 2 (agent type: AI vs. human) \times 2 (hiding strategy: evasive vs. rationalized) between-subjects design. A total of 212 participants (119 women, $M_{\text{age}} = 30.41$, $SD = 10.51$) were recruited from Credamo.

Procedure.

Participants were randomly assigned to one of the four experimental conditions. They were asked to imagine using a popular social media app and discovering that some of the posts and comments they had previously made had disappeared. They then reported this issue through the app's user feedback function. Afterward, they received a response from a service agent. The manipulation of agent type was identical to Study 1. In the evasive hiding conditions, the agent replied: "Hello, we are aware of this issue. I have forwarded your issue to our technical team for further investigation and aim to fix it in the future. Thank you for your understanding." In the rationalized hiding conditions, the agent replied: "Hello, due to recent uncontrollable network turbulence, some user data was lost. Besides our app, many other apps are also affected. Thank you for your understanding."

After reading the materials, participants indicated the agent type and rated the hiding strategies. They then reported their willingness to forgive. To test the

underlying processes, participants also reported their perceptions of the agent's response, specifically perceived negative motives (Joireman *et al.*, 2013) and perceived sincerity (Barasch *et al.*, 2016). Prior research suggested that user responses may be influenced by the clarity and credibility of the agent's reply (Antonetti and Baghi, 2023), as well as by users' inferences about the agents' selfish and benevolent intentions from the agent's communication (Garvey *et al.*, 2023). These variables were measured to rule out alternative explanations. Finally, participants provided their demographic information. The full measurement items were provided in Appendix B.

Results

Manipulation check.

The results indicated that participants in the evasive hiding conditions scored significantly higher on the evasive hiding measure than those in the rationalized hiding conditions ($M_{\text{Evasive}} = 4.43$, $SD = 1.68$; $M_{\text{Rationalized}} = 3.26$, $SD = 1.41$; $F(1, 210) = 30.12$, $p < 0.001$), while participants in the rationalized hiding conditions scored significantly higher on the rationalized hiding measure than those in the evasive hiding conditions ($M_{\text{Rationalized}} = 5.65$, $SD = 1.18$; $M_{\text{Evasive}} = 3.39$, $SD = 1.77$; $F(1, 210) = 120.25$, $p < 0.001$).

Willingness to forgive.

The results of a 2 (agent type: AI vs. human) \times 2 (hiding strategy: evasive vs. rationalized) ANOVA revealed a significant interaction effect ($F(1, 208) = 11.97$, $p < 0.01$) and a significant main effect of hiding strategy ($F(1, 208) = 19.77$, $p < 0.001$). The main effect of agent type was not significant ($F(1, 208) = 0.02$, $p > 0.05$). A simple effect analysis showed that participants demonstrated a higher willingness to forgive when the evasive hiding strategy was employed by AI agents rather than human agents ($M_{\text{AI}} = 5.54$, $SD = 1.70$; $M_{\text{Human}} = 4.80$, $SD = 1.64$; $F(1, 208) = 5.46$, $p < 0.05$). Conversely, participants showed a higher willingness to forgive when the rationalized hiding strategy was employed by human agents rather than AI agents ($M_{\text{Human}} = 4.58$, $SD = 1.47$; $M_{\text{AI}} = 3.77$, $SD = 1.66$; $F(1, 208) = 6.54$, $p < 0.05$) (see

Figure 3).

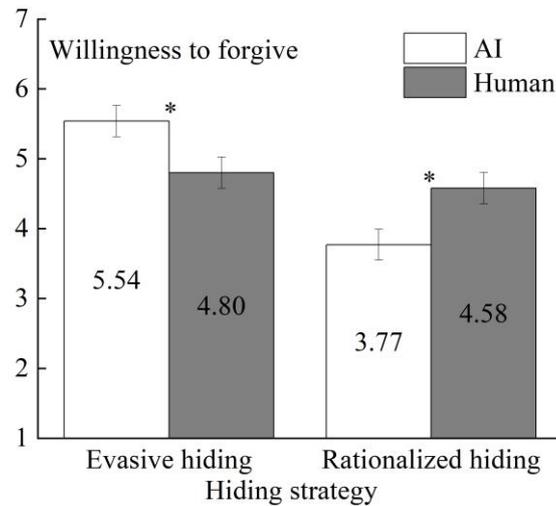


Figure 3. Study 2: Willingness to forgive as a function of agent type and hiding strategy

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; Error bars = ± 1 SEs.

Source(s): Authors' own work

Perceived negative motives and perceived sincerity.

Before the moderated mediation analysis, the effects of agent type and hiding strategy on the psychological mediators (perceived negative motives and perceived sincerity) were examined. A 2 (agent type: AI vs. human) \times 2 (hiding strategy: evasive vs. rationalized) ANOVA was conducted on perceived negative motives, revealing a significant main effect of hiding strategy ($F(1, 208) = 9.72, p < 0.01$) and a significant interaction effect ($F(1, 208) = 6.53, p < 0.05$), while the main effect of agent type was not significant ($p > 0.05$). A simple effect analysis revealed that, under the evasive hiding strategy, perceived negative motives were significantly lower in the AI agent condition than in the human agent condition ($M_{AI} = 2.69, SD = 1.44; M_{Human} = 3.32, SD = 1.70; F(1, 208) = 4.17, p < 0.05$). By contrast, under the rationalized hiding strategy, perceived negative motives did not differ significantly between the AI agent condition and the human agent condition ($M_{AI} = 3.92, SD = 1.49; M_{Human} = 3.44, SD = 1.65; F(1, 208) = 2.47, p > 0.05$).

Similarly, a 2×2 ANOVA of perceived sincerity revealed a significant main effect of hiding strategy ($F(1, 208) = 4.95, p < 0.05$) and a significant interaction effect ($F(1, 208) = 6.95, p < 0.01$), while the main effect of agent type was not significant ($p > 0.05$). A simple effect analysis revealed that, under the rationalized hiding strategy, perceived sincerity was significantly higher in the human agent condition than in the AI agent condition ($M_{\text{Human}} = 4.60, SD = 1.54; M_{\text{AI}} = 3.93, SD = 1.60; F(1, 208) = 5.00, p < 0.05$). However, under the evasive hiding strategy, perceived sincerity did not differ significantly between the AI agent condition and the human agent condition ($M_{\text{AI}} = 4.02, SD = 1.56; M_{\text{Human}} = 3.58, SD = 1.40; F(1, 208) = 2.23, p > 0.05$).

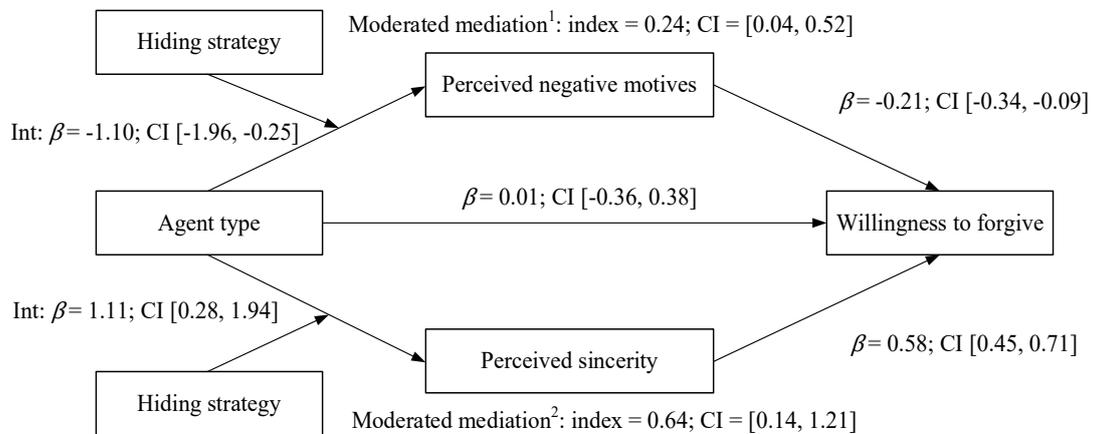
Moderated mediation.

We conducted a moderated mediation analysis using Model 7 of the PROCESS macro for SPSS (Hayes, 2017). Agent type was the dichotomous independent variable (0 = AI, 1 = human), hiding strategy were the dichotomous moderator (0 = evasive hiding, 1 = rationalized hiding), perceived negative motives and perceived sincerity were the mediators, and willingness to forgive was the dependent variable. The specific results were shown in Table 2. A bootstrapping analysis with 5,000 samples revealed significant moderated mediation effects for perceived negative motives (index of moderated mediation = 0.24, 95% CI = [0.04, 0.52]). Specifically, in the evasive hiding condition, agent type significantly influenced willingness to forgive via perceived negative motives, supporting the mediating role of perceived negative motives ($\beta = -0.13, 95\% \text{ CI} = [-0.34, -0.01]$). However, this indirect effect was not significant in the rationalized hiding condition ($\beta = 0.10, 95\% \text{ CI} = [-0.03, 0.27]$). As shown in Figure 4, the results indicated that perceived negative motives mediated the effect of agent type on users' willingness to forgive when an evasive hiding strategy was used. The results supported H2a.

[Insert Table 2]

Similarly, this analysis revealed significant moderated mediation effects for

perceived sincerity (index of moderated mediation = 0.64, 95% CI = [0.14, 1.21]). Specifically, in the rationalized hiding condition, agent type influenced willingness to forgive through perceived sincerity ($\beta = 0.38$, 95% CI = [0.03, 0.77]), supporting the mediating role of perceived sincerity. However, in the evasive hiding condition, the indirect effect was not significant ($\beta = -0.26$, 95% CI = [-0.62, 0.07]). The results revealed that perceived sincerity mediated the effect of agent type on users' willingness to forgive when a rationalized hiding strategy was employed. The results supported H2b.



Int: Interaction term (*Agent type* × *Hiding strategy*)
 Moderated mediation¹: *Perceived negative motives* as a mediator
 Moderated mediation²: *Perceived sincerity* as a mediator

Figure 4. Moderated mediation analysis
Source(s): Authors' own work

The ANOVA results indicated no interaction effects between agent type and hiding strategy on clarity ($F(1, 208) = 1.28, p > 0.05$), credibility ($F(1, 208) = 1.74, p > 0.05$), benevolent intentions ($F(1, 208) = 1.09, p > 0.05$), or selfish intentions ($F(1, 208) = 2.78, p > 0.05$). Taken together, these findings served to rule out the alternative explanations for these variables and enhance the robustness of the conceptual model.

Study 2 elucidated the psychological processes underlying the interaction effects of agent type and hiding strategy. Specifically, in the evasive hiding condition, AI agents were perceived to possess lower negative motives than human agents, which in

turn enhanced users' willingness to forgive. In contrast, in the rationalized hiding condition, human agents were perceived as more sincere than AI agents, thereby increasing users' willingness to forgive. In addition, these analyses ruled out multiple potential alternative explanations.

Study 3

Study 3 shifted the focus from the human-AI comparison to different types of AI agents and examined how AI capabilities (mechanical AI vs. thinking AI) influence user responses to hiding strategies. Following previous studies, Study 3 was conducted in a new, tricky complaint scenario regarding excessive app advertising notifications.

Method

Participants and design.

Study 3 employed a 2 (agent type: mechanical AI vs. thinking AI) \times 2 (hiding strategy: evasive vs. rationalized) between-subjects design. A total of 302 participants (155 women, $M_{\text{age}} = 34.37$, $SD = 8.05$) were recruited.

Procedure.

Participants were randomly assigned to one of the four experimental conditions. They were asked to imagine using a popular social app and encountering an increasing number of pop-up advertisements when opening the app. In addition, the app occasionally redirected users automatically to ad pages or other mobile apps; consequently, they reported this issue to the user service function. Participants were informed that their feedback was processed by an AI agent and read a description of either a mechanical or a thinking AI agent. The manipulation of the two types of AI agents was conveyed through descriptions of their conversational and comprehension capabilities, such as whether the agent could handle only basic dialogue scenarios or could understand contextual cues and generate coherent responses similar to those

produced by humans, rather than technical terms such as “natural language processing” or “deep learning”. This manipulation approach was adapted from Srinivasan and Sarial-Abi (2021).

Subsequently, they received a reply from the agent. In the evasive hiding conditions, the agent replied: “Hello, your feedback has been forwarded to relevant teams for evaluation and processing. Thank you for your understanding.” In the rationalized hiding conditions, the agent replied: “Hello, the function has been verified and approved by the relevant teams. We recommend holding your phone stable to avoid activation of the shaking function when opening the app. Thank you for your understanding.” Participants were then asked to evaluate the agent’s capability (Garvey *et al.*, 2023). To examine whether the underlying mechanisms identified in Study 2 also apply to the comparison between the two types of AI agents, participants reported their perceptions of the agent’s negative motives and sincerity. In addition, prior research suggested that manipulations of AI appearance or capability may influence user behavior through perceived social presence (Konya-Baumbach *et al.*, 2023; Munnukka *et al.*, 2022). To rule out this alternative explanation, participants also evaluated the agent’s social presence (Verhagen *et al.*, 2014). The measurement items were provided in Appendix B.

Results

Manipulation check.

The results indicated that participants in the thinking AI conditions scored significantly higher on AI capability than those in the mechanical AI conditions ($M_{\text{Thinking AI}} = 5.34$, $SD = 0.85$; $M_{\text{Mechanical AI}} = 4.49$, $SD = 1.30$; $F(1, 300) = 44.66$, $p < 0.001$), confirming the successful manipulation of the type of AI agent. In addition, participants in the evasive hiding conditions scored significantly higher on the evasive hiding measure than those in the rationalized hiding conditions ($M_{\text{Evasive}} = 4.80$, $SD = 1.78$; $M_{\text{Rationalized}} = 3.65$, $SD = 1.52$; $F(1, 300) = 36.40$, $p < 0.001$), while participants in the rationalized hiding conditions scored significantly higher on the rationalized

hiding measure than those in the evasive hiding conditions ($M_{\text{Rationalized}} = 4.93$, $SD = 1.42$; $M_{\text{Evasive}} = 3.48$, $SD = 1.65$; $F(1, 300) = 66.91$, $p < 0.001$), thereby confirming the successful manipulation of the hiding strategies.

Willingness to forgive.

The results of a 2 (agent type: mechanical AI vs. thinking AI) \times 2 (hiding strategy: evasive vs. rationalized) ANOVA revealed a significant interaction effect ($F(1, 298) = 27.72$, $p < 0.001$). The main effects of agent type ($F(1, 298) = 0.01$, $p > 0.05$) and hiding strategy ($F(1, 298) = 0.75$, $p > 0.05$) were not significant. A simple effect analysis showed that participants demonstrated a higher willingness to forgive when the evasive hiding strategy was employed by mechanical AI agents compared to thinking AI agents ($M_{\text{Mechanical AI}} = 4.17$, $SD = 1.64$; $M_{\text{Thinking AI}} = 3.22$, $SD = 1.50$; $F(1, 298) = 13.39$, $p < 0.001$). Conversely, participants showed a higher willingness to forgive when the rationalized hiding strategy was employed by thinking AI agents compared to mechanical AI agents ($M_{\text{Thinking AI}} = 4.03$, $SD = 1.59$; $M_{\text{Mechanical AI}} = 3.04$, $SD = 1.65$; $F(1, 298) = 14.34$, $p < 0.001$) (see Figure 5). Thus, H3a and H3b were supported.

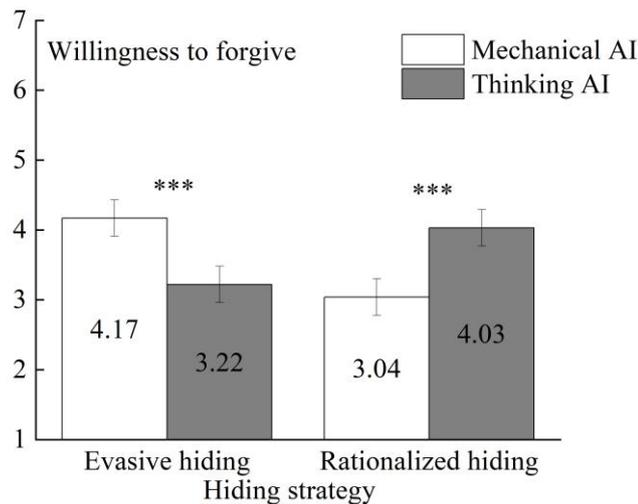


Figure 5. Study 3: Willingness to forgive as a function of agent type and hiding strategy

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; Error bars = ± 1 SEs.

Source(s): Authors' own work

This study further examined whether perceived negative motives and perceived sincerity could account for the interaction effect between agent type (mechanical AI vs. thinking AI) and hiding strategy on users' willingness to forgive. The 2×2 ANOVA on perceived negative motives revealed a significant main effect of hiding strategy ($F(1, 298) = 5.84, p < 0.05$). Neither the main effect of AI type ($F(1, 298) = 1.90, p > 0.05$) nor the interaction effect ($F(1, 298) = 0.15, p > 0.05$) was significant. Similarly, the 2×2 ANOVA on perceived sincerity revealed a significant main effect of hiding strategy ($F(1, 298) = 4.91, p < 0.05$). Neither the main effect of AI type ($F(1, 298) = 0.77, p > 0.05$) nor the interaction effect ($F(1, 298) = 0.29, p > 0.05$) was significant. In addition, the ANOVA results for social presence indicated a non-significant interaction effect between AI type and hiding strategy ($F(1, 298) = 0.38, p > 0.05$).

The results of Study 3 provided support for H3a and H3b, demonstrating that when encountering an evasive hiding strategy, participants exhibited a higher willingness to forgive mechanical AI agents; conversely, when encountering a rationalized hiding strategy, they showed a higher willingness to forgive thinking AI agents. This finding indicated that thinking AI, owing to its enhanced capabilities, elicited user responses similar to those observed in human agent conditions. However, further analysis ruled out the mediating roles of perceived negative motives and perceived sincerity in the interaction effect between AI type and hiding strategy. Social presence, as an alternative explanation, was also excluded.

General discussion

This research focuses on the interaction effect between agent type and intention hiding strategy in the context of organizational responses to tricky user complaints. The results indicated that users exhibited a higher willingness to forgive when an evasive hiding strategy was used by an AI agent or when a rationalized hiding strategy

was used by a human agent. Such alignment between agent type and intention hiding strategy can effectively enhance user forgiveness. This observation aligns with prior research suggesting that individuals may respond to AI and human agents in divergent ways depending on the nature of the service context (Garvey *et al.*, 2023; Longoni and Cian, 2022; Xu and Mehta, 2022).

Furthermore, this research uncovers the underlying psychological mechanisms that explain this interaction effect. According to the service recovery literature, individuals tend to exhibit greater tolerance toward service failure when they perceive the company or its representatives as more sincere or as acting with positive intentions (Antonetti and Baghi, 2023; Béal and Grégoire, 2022). The results supported this notion and demonstrated that the effect of agent type and intention hiding strategy on willingness to forgive was mediated by perceived negative motives and perceived sincerity. Specifically, when an AI agent used evasive hiding, users perceived a lower level of negative motives, thereby increasing their willingness to forgive. Conversely, when a human agent used rationalized hiding, users perceived greater sincerity, which in turn enhanced their forgiveness. These findings are consistent with mind perception theory, which posits that individuals evaluate both human and nonhuman agents based on perceived agency and experience, and that such evaluations shape their subsequent attitudes and behaviors (Chen *et al.*, 2024b; Gray *et al.*, 2007; Lan *et al.*, 2025).

This research further reveals an interaction effect between AI capability and intention hiding strategy. Specifically, users reported a higher willingness to forgive when mechanical AI employed an evasive hiding strategy or when thinking AI adopted a rationalized hiding strategy. This finding provides important implications in two respects. First, it responds to the call by Huang and Rust (2021) to examine how AI agents with different levels of capabilities influence user responses in more complex service contexts. Second, it supports the argument made by Garvey *et al.* (2023), suggesting that tailoring AI agent characteristics can elicit cognitive and

emotional reactions from users that are similar to those triggered by human agents.

Theoretical implications

This research makes several contributions to existing literature. First, it integrates the intention hiding framework (Connelly *et al.*, 2012) into the service recovery literature by defining evasive and rationalized hiding as strategic linguistic tactics for managing tricky complaints under resource constraints. Positioning these strategies along dimensions of information transparency and responsibility acceptance, the research complements traditional recovery approaches such as apology and compensation (Antonetti and Baghi, 2023; Grégoire *et al.*, 2025). Furthermore, by examining AI as the executor of these strategies, the research explores the interaction between agent types and strategy types. This addresses a significant gap regarding how executor characteristics influence user responses and offers an integrative theoretical framework for understanding the psychological processes when AI is involved in complex service dialogues.

Second, this research introduces mind perception theory into the context of AI agents employing intention hiding strategies and validates its robustness in the service recovery domain. Prior applications of mind perception theory have primarily focused on rational tasks such as product recommendation and risk assessment, with an emphasis on how users respond to AI in cognitive decisions (Jin *et al.*, 2025; Lan *et al.*, 2025; Srinivasan and Sarial-Abi, 2021). Relatively little research has addressed emotionally charged and relationship-oriented service recovery contexts. By revealing users' dual perceptions of agent intentionality and emotional capacity, this research offers a new perspective for explaining users' psychological judgments when encountering different combinations of agent type and intention hiding strategy. In doing so, it extends the application of mind perception theory to service recovery contexts that involve greater social complexity.

Third, this research incorporates AI capability (mechanical AI vs. thinking AI)

into the theoretical framework of service recovery, thereby deepening our understanding of how technical attributes influence users' social cognition. Mechanical and thinking AI represent two functionally distinct types of intelligent agents that have attracted growing attention in customer interaction research (Gursoy and Cai, 2025). However, most existing studies have focused on their role in task alignment or service performance, while limited discussion is found in emotional interactions and service recovery contexts (Lan *et al.*, 2025; Longoni and Cian, 2022). Highlighting AI capability as a manageable and configurable feature of AI agents, this research shows the importance of technical attributes in shaping users' interpretation and psychological processes. It contributes to the literature on how digital technologies facilitate relationship recovery, emotional regulation, and service communication. It also opens new directions for future research on how internal variations among AI agents influence user trust, attribution, and judgment.

Managerial implications

To effectively manage tricky complaints, firms should align intention hiding strategies with agent type to minimize negative user attributions. Managers are advised to prioritize evasive hiding for AI agents by utilizing structured and ambiguous responses such as “the issue has been forwarded to the relevant team”. Conversely, rationalized hiding strategies that invoke institutional policies or external constraints should be primarily delivered by human agents. This strategic differentiation leverages the reduced perceived agency of AI to mitigate negative motive attributions while utilizing human credibility and social presence to enhance the willingness to forgive.

Organizations should further align intention hiding strategies with AI capability by pairing thinking AI with rationalized hiding and mechanical AI with evasive hiding. This alignment must be supported by clear AI labels that communicate the AI agent's capability boundaries, thereby helping users form appropriate expectations. Furthermore, companies should implement a dynamic monitoring system that tracks

metrics such as complaint escalation rates and satisfaction scores to facilitate real-time adjustments. Such a system allows for immediate intervention, including switching to a human agent or modifying response phrasing, when a strategy-agent mismatch is detected in high-pressure service contexts.

Limitations and future directions

We acknowledge limitations in the present research. From a theoretical perspective, the research primarily focuses on service agent types and hiding strategies and does not consider user characteristics or firm attributes. Prior studies have demonstrated that these factors can substantially shape user preferences for and responses to AI-delivered services (Logg *et al.*, 2019; Xu and Mehta, 2022). Future research could further examine the potential effects of these factors on our findings.

Notably, the results of Study 3 indicated that the mediating roles of perceived negative motives and perceived sincerity were not supported in the interaction between AI types (mechanical vs. thinking AI) and hiding strategies. These findings suggest that in purely AI contexts, users rely heavily on evaluations of whether AI capabilities align with the employed strategies, rather than on attributions of motives and emotions. Consistency between hiding strategies and the presumed capabilities of AI is more likely to foster expectation consistency and lenient judgments among users (Crolig *et al.*, 2022). Future research may validate this explanation centered on cognitive expectation confirmation (Laato *et al.*, 2022; Qi *et al.*, 2025; Wu *et al.*, 2024).

In addition, hiding strategies themselves entail potential ethical concerns. Given that complaint handling is often regarded as a moral obligation that firms owe to aggrieved users, non-substantive responses may trigger moral concerns and lead to adverse consequences (Honora *et al.*, 2025). Future research could examine users' assessments of the moral legitimacy of evasive and rationalized hiding strategies from ethical perspectives. Additionally, it could investigate whether perceived moral

transgressions, when such strategies are interpreted as communicative deception or violations of benevolence norms, undermine service recovery effectiveness through pathways such as negative emotional responses.

While the experimental methods in this study ensure high internal validity, future research could incorporate field experiments or behavioral data to enhance ecological validity. Although this study measures users' immediate willingness to forgive, longitudinal measures could be employed to examine the effectiveness of hiding strategies over time.

References

- Anand, A., Offergelt, F. and Anand, P. (2022), "Knowledge hiding: a systematic review and research agenda", *Journal of Knowledge Management*, Vol. 26 No. 6, pp. 1438-1457, doi: 10.1108/JKM-04-2021-0336.
- Antonetti, P. and Baghi, I. (2023), "Projecting lower competence to boost apology effectiveness: underlying mechanism and boundary conditions", *Journal of the Academy of Marketing Science*, Vol. 51 No. 3, pp. 695-715, doi: 10.1007/s11747-022-00903-5.
- Barasch, A., Berman, J. Z. and Small, D. A. (2016), "When payment undermines the pitch: on the persuasiveness of pure motives in fund-raising", *Psychological Science*, Vol. 27 No. 10, pp. 1388-1397, doi: 10.1177/0956797616638841.
- Béal, M. and Grégoire, Y. (2022), "How do observers react to companies' humorous responses to online public complaints?", *Journal of Service Research*, Vol. 25 No. 2, pp. 242-259, doi: 10.1177/1094670521989448.
- Bergner, A. S., Hildebrand, C. and Häubl, G. (2023), "Machine talk: how verbal embodiment in conversational AI shapes consumer-brand relationships", *Journal of Consumer Research*, Vol. 50 No. 4, pp. 742-764, doi: 10.1093/jcr/ucad014.
- Candrian, C. and Scherer, A. (2022), "Rise of the machines: delegating decisions to autonomous AI", *Computers in Human Behavior*, Vol. 134, p. 107308, doi: 10.1016/j.chb.2022.107308.
- Chaudhry, S. J. and Loewenstein, G. (2019), "Thanking, apologizing, bragging, and blaming: responsibility exchange theory and the currency of communication", *Psychological Review*, Vol. 126 No. 3, pp. 313-344, doi: 10.1037/rev0000139.
- Chen, L., Zhang, S., Hu, X., Liu, S. and Lan, R. (2024a), "Does political skill promote knowledge hiding? An impression management perspective", *Journal of Knowledge Management*, Vol. 28 No. 6, pp. 1632-1647, doi: 10.1108/JKM-05-2023-0431.
- Chen, Q., Gong, Y., Lu, Y. and Luo, X. (2025), "The golden zone of AI's emotional expression in frontline chatbot service failures", *Internet Research*, Vol. 35 No. 3,

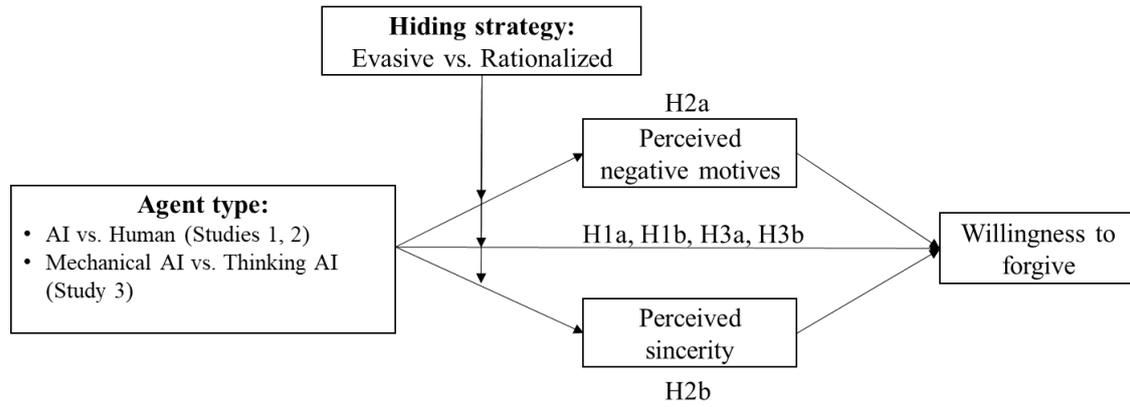
-
- pp. 1065-1103, doi: 10.1108/INTR-07-2023-0551.
- Chen, Y., Wang, H., Hill, S. R. and Li, B. (2024b), "Consumer attitudes toward AI-generated ads: appeal types, self-efficacy and AI's social role", *Journal of Business Research*, Vol. 185, p. 114867, doi: 10.1016/j.jbusres.2024.114867.
- Cheng, L. K. (2023), "Effects of service robots' anthropomorphism on consumers' attribution toward and forgiveness of service failure", *Journal of Consumer Behaviour*, Vol. 22 No. 1, pp. 67-81, doi: 10.1002/cb.2112.
- Cheng, X., Qiao, L., Yang, B. and Li, Z. (2024), "An investigation on the influencing factors of elderly people's intention to use financial AI customer service", *Internet Research*, Vol. 34 No. 3, pp. 690-717, doi: 10.1108/INTR-06-2022-0402.
- Choi, S., Mattila, A. S. and Bolton, L. E. (2021), "To err is human (-oid): how do consumers react to robot service failure and recovery?", *Journal of Service Research*, Vol. 24 No. 3, pp. 354-371, doi: 10.1177/1094670520978798.
- Connelly, C. E., Zweig, D., Webster, J. and Trougakos, J. P. (2012), "Knowledge hiding in organizations", *Journal of Organizational Behavior*, Vol. 33 No. 1, pp. 64-88, doi: 10.1002/job.737.
- Crolic, C., Thomaz, F., Hadi, R. and Stephen, A. T. (2022), "Blame the bot: anthropomorphism and anger in customer-chatbot interactions", *Journal of Marketing*, Vol. 86 No. 1, pp. 132-148, doi: 10.1177/00222429211045687.
- Fan, A., Wu, L., Miao, L. and Mattila, A. S. (2020), "When does technology anthropomorphism help alleviate customer dissatisfaction after a service failure? The moderating role of consumer technology self-efficacy and interdependent self-construal", *Journal of Hospitality Marketing & Management*, Vol. 29 No. 3, pp. 269-290, doi: 10.1080/19368623.2019.1639095.
- Garvey, A. M., Kim, T. and Duhachek, A. (2023), "Bad news? Send an AI. Good news? Send a human", *Journal of Marketing*, Vol. 87 No. 1, pp. 10-25, doi: 10.1177/00222429211066972.
- Gong, Y., Tang, X., Chang, E.-C. and Wan, F. (2025), "Corporate sociopolitical activism and language strategies for mitigating animosity in a host country", *Journal of International Marketing*, Vol. 33 No. 3-4, pp. 28-43, doi: 10.1177/1069031X251319018.
- Gray, H. M., Gray, K. and Wegner, D. M. (2007), "Dimensions of mind perception", *Science*, Vol. 315 No. 5812, pp. 619-619, doi: 10.1126/science.1134475.
- Grégoire, Y., Khamitov, M., Carrillat, F. A. and Rohani, M. (2025), "The attenuation effects of time and "sensemaking" surveys on customer revenge", *Journal of the Academy of Marketing Science*, Vol. 53 No. 1, pp. 172-196, doi: 10.1007/s11747-024-01046-5.
- Güntürkün, P., Haumann, T. and Mikolon, S. (2020), "Disentangling the differential roles of warmth and competence judgments in customer-service provider relationships", *Journal of Service Research*, Vol. 23 No. 4, pp. 476-503, doi: 10.1177/1094670520920354.
- Gursoy, D. and Cai, R. (2025), "Artificial intelligence: an overview of research trends and future directions", *International Journal of Contemporary Hospitality*

-
- Management*, Vol. 37 No. 1, pp. 1-17, doi: 10.1108/IJCHM-03-2024-0322.
- Han, E., Yin, D. and Zhang, H. (2023), "Bots with feelings: should AI agents express positive emotion in customer service?", *Information Systems Research*, Vol. 34 No. 3, pp. 1296-1311, doi: 10.1287/isre.2022.1179.
- Hayes, A. F. (2017), *Introduction to Mediation, Moderation, and Conditional Process Analysis: a Regression-based Approach*, Guilford Publications, New York.
- Holthöwer, J. and Van Doorn, J. (2023), "Robots do not judge: service robots can alleviate embarrassment in service encounters", *Journal of the Academy of Marketing Science*, Vol. 51 No. 4, pp. 767-784, doi: 10.1007/s11747-022-00862-x.
- Honora, A., Japutra, A. and Septianto, F. (2025), "Don't humor me! Customers' moral perceptions toward companies' humorous responses in social media service recovery", *Journal of Business Ethics*, pp. 1-19, doi: 10.1007/s10551-025-06154-y.
- Huang, M. H. and Rust, R. T. (2021), "Engaged to a robot? The role of AI in service", *Journal of Service Research*, Vol. 24 No. 1, pp. 30-41, doi: 10.1177/1094670520902266.
- Javornik, A., Filieri, R. and Gumann, R. (2020), "Don't forget that others are watching, too! The effect of conversational human voice and reply length on observers' perceptions of complaint handling in social media", *Journal of Interactive Marketing*, Vol. 50 No. 1, pp. 100-119, doi: 10.1016/j.intmar.2020.02.002.
- Jin, J., Walker, J. and Reczek, R. W. (2025), "Avoiding embarrassment online: response to and inferences about chatbots when purchases activate self-presentation concerns", *Journal of Consumer Psychology*, Vol. 35 No. 2, pp. 185-202, doi: 10.1002/jcpy.1414.
- Joireman, J., Grégoire, Y., Devezer, B. and Tripp, T. M. (2013), "When do customers offer firms a "second chance" following a double deviation? The impact of inferred firm motives on customer revenge and reconciliation", *Journal of Retailing*, Vol. 89 No. 3, pp. 315-337, doi: 10.1016/j.jretai.2013.03.002.
- Kim, T. W., Jiang, L., Duhachek, A., Lee, H. and Garvey, A. (2022), "Do you mind if I ask you a personal question? How AI service agents alter consumer self-disclosure", *Journal of Service Research*, Vol. 25 No. 4, pp. 649-666, doi: 10.1177/10946705221120232.
- Konya-Baumbach, E., Biller, M. and von Janda, S. (2023), "Someone out there? A study on the social presence of anthropomorphized chatbots", *Computers in Human Behavior*, Vol. 139, p. 107513, doi: 10.1016/j.chb.2022.107513.
- Laato, S., Tiainen, M., Najmul Islam, A. K. M. and Mäntymäki, M. (2022), "How to explain AI systems to end users: a systematic literature review and research agenda", *Internet Research*, Vol. 32 No. 7, pp. 1-31, doi: 10.1108/INTR-08-2021-0600.
- Lan, H., Tang, X., Fu, M. and Peng, H. (2025), "Will human designers be replaced? Exploring consumer responses to AI involvement in interior design", *Current Psychology*, Vol. 44 No. 9, pp. 8146-8157, doi: 10.1007/s12144-025-07422-w.
- Liu, D., Lv, Y. and Huang, W. (2023), "How do consumers react to chatbots'

-
- humorous emojis in service failures”, *Technology in Society*, Vol. 73, p. 102244, doi: 10.1016/j.techsoc.2023.102244.
- Liu, J. and Xu, X. A. (2023), “Humor type and service context shape AI service recovery”, *Annals of Tourism Research*, Vol. 103, p. 103668, doi: 10.1016/j.annals.2023.103668.
- Logg, J. M., Minson, J. A. and Moore, D. A. (2019), “Algorithm appreciation: people prefer algorithmic to human judgment”, *Organizational Behavior and Human Decision Processes*, Vol. 151, pp. 90-103, doi: 10.1016/j.obhdp.2018.12.005.
- Longoni, C. and Cian, L. (2022), “Artificial intelligence in utilitarian vs. hedonic contexts: the “word-of-machine” effect”, *Journal of Marketing*, Vol. 86 No. 1, pp. 91-108, doi: 10.1177/0022242920957347.
- Longoni, C., Cian, L. and Kyung, E. J. (2023), “Algorithmic transference: people overgeneralize failures of AI in the government”, *Journal of Marketing Research*, Vol. 60 No. 1, pp. 170-188, doi: 10.1177/00222437221110139.
- Lv, X., Liu, Y., Luo, J., Liu, Y. and 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*, Vol. 87, p. 103114, doi: 10.1016/j.annals.2020.103114.
- Lv, X., Yang, Y., Qin, D., Cao, X. and Xu, H. (2022), “Artificial intelligence service recovery: the role of empathic response in hospitality customers’ continuous usage intention”, *Computers in Human Behavior*, Vol. 126, p. 106993, doi: doi.org/10.1016/j.chb.2021.106993.
- Michalová, T., Maršíková, K., Falát, L. and Madzík, P. (2024), “Altruistic leadership and its role in reducing knowledge hiding: the mediating effects of team learning and knowledge culture”, *Journal of Innovation & Knowledge*, Vol. 9 No. 4, p. 100592, doi: 10.1016/j.jik.2024.100592.
- Munnukka, J., Talvitie-Lamberg, K. and Maity, D. (2022), “Anthropomorphism and social presence in human-virtual service assistant interactions: the role of dialog length and attitudes”, *Computers in Human Behavior*, Vol. 135, p. 107343, doi: 10.1016/j.chb.2022.107343.
- Muzatko, S. and Bansal, G. (2024), “It pays to be forthcoming: timing of data breach announcement, trust violation, and trust restoration”, *Internet Research*, Vol. 34 No. 5, pp. 1629-1663, doi: 10.1108/INTR-12-2021-0939.
- Ohtsubo, Y. and Watanabe, E. (2009), “Do sincere apologies need to be costly? Test of a costly signaling model of apology”, *Evolution and Human Behavior*, Vol. 30 No. 2, pp. 114-123, doi: 10.1016/j.evolhumbehav.2008.09.004.
- Peter, S., Riemer, K. and West, J. D. (2025), “The benefits and dangers of anthropomorphic conversational agents”, *Proceedings of the National Academy of Sciences*, Vol. 122 No. 22, p. e2415898122, doi: 10.1073/pnas.2415898122.
- Pitardi, V., Wirtz, J., Paluch, S. and Kunz, W. H. (2024), “Metaperception benefits of service robots in uncomfortable service encounters”, *Tourism Management*, Vol. 105, p. 104939, doi: 10.1016/j.tourman.2024.104939.
- Qi, T., Liu, H. and Huang, Z. (2025), “An assistant or a friend? The role of parasocial relationship of human-computer interaction”, *Computers in Human Behavior*,

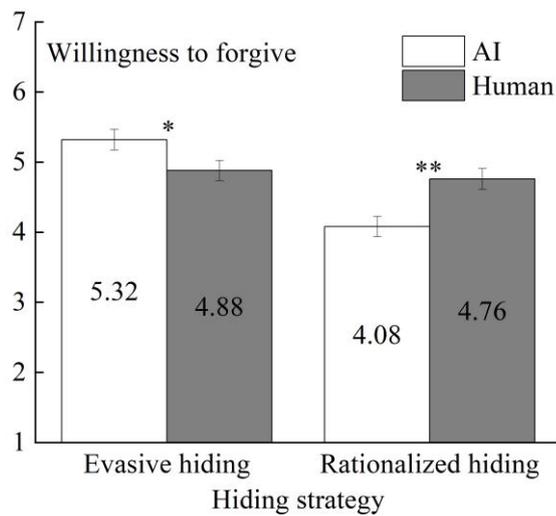
-
- Vol. 167, p. 108625, doi: doi.org/10.1016/j.chb.2025.108625.
- Schepers, J., Belanche, D., Casaló, L. V. and Flavián, C. (2022), “How smart should a service robot be?”, *Journal of Service Research*, Vol. 25 No. 4, pp. 565-582, doi: 10.1177/10946705221107704.
- Srinivasan, R. and Sarial-Abi, G. (2021), “When algorithms fail: consumers’ responses to brand harm crises caused by algorithm errors”, *Journal of Marketing*, Vol. 85 No. 5, pp. 74-91, doi: 10.1177/0022242921997082.
- Tang, X., Chang, E. C., Huang, X. and Zhang, M. (2018), “Timing and compensation strategies in service recovery”, *Journal of Services Marketing*, Vol. 32 No. 6, pp. 755-766, doi: 10.1108/JSM-04-2017-0126.
- Verhagen, T., Van Nes, J., Feldberg, F. and Van Dolen, W. (2014), “Virtual customer service agents: using social presence and personalization to shape online service encounters”, *Journal of Computer-Mediated Communication*, Vol. 19 No. 3, pp. 529-545, doi: 10.1111/jcc4.12066.
- Wang, K. Y., Chih, W. H. and Honora, A. (2023), “How the emoji use in apology messages influences customers’ responses in online service recoveries: the moderating role of communication style”, *International Journal of Information Management*, Vol. 69, p. 102618, doi: 10.1016/j.ijinfomgt.2022.102618.
- Wei, H. and Ran, Y. (2019), “Male versus female: how the gender of apologizers influences consumer forgiveness”, *Journal of Business Ethics*, Vol. 154 No. 2, pp. 371-387, doi: 10.1007/s10551-017-3440-7.
- Wu, X., Zhou, Z. and Chen, S. (2024), “A mixed-methods investigation of the factors affecting the use of facial recognition as a threatening AI application”, *Internet Research*, Vol. 34 No. 5, pp. 1872-1897, doi: 10.1108/INTR-11-2022-0894.
- Xie, Y., Tong, Z. and Wu, Z. (2025a), “Artificial intelligence or human service, which customer service failure is more unforgivable? A counterfactual thinking perspective”, *Psychology & Marketing*, Vol. 42 No. 8, pp. 1993-2011, doi: 10.1002/mar.22215.
- Xie, Y., Zhou, P., Liang, C., Zhao, S. and Lu, W. (2025b), “Affiliative or self-defeating? Exploring the effect of humor types on customer forgiveness in the context of AI agents’ service failure”, *Journal of Business Research*, Vol. 194, p. 115381, doi: 10.1016/j.jbusres.2025.115381.
- Xu, L. and Mehta, R. (2022), “Technology devalues luxury? Exploring consumer responses to AI-designed luxury products”, *Journal of the Academy of Marketing Science*, Vol. 50 No. 6, pp. 1135-1152, doi: 10.1007/s11747-022-00854-x.
- Yalcin, G., Lim, S., Puntoni, S. and van Osselaer, S. M. (2022), “Thumbs up or down: consumer reactions to decisions by algorithms versus humans”, *Journal of Marketing Research*, Vol. 59 No. 4, pp. 696-717, doi: 10.1177/00222437211070016.
- Zhao, H., Jiang, L. and Su, C. (2020), “To defend or not to defend? How responses to negative customer review affect prospective customers’ distrust and purchase intention”, *Journal of Interactive Marketing*, Vol. 50 No. 1, pp. 45-64, doi: 10.1016/j.intmar.2019.11.001.

Plate 1: The research model



Source(s): Authors' own work

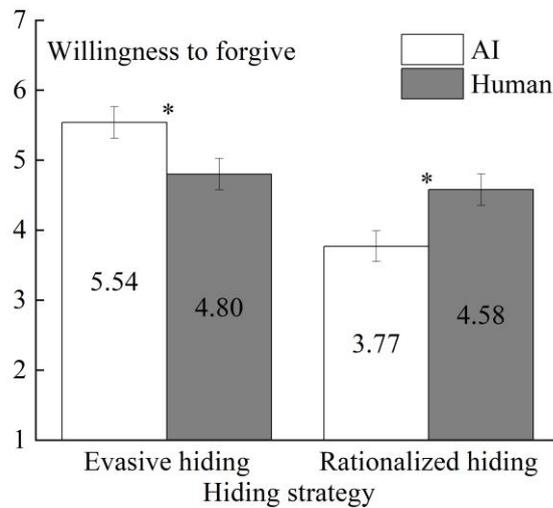
Plate 2: Study 1: Willingness to forgive as a function of agent type and hiding strategy



Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; Error bars = ± 1 SEs.

Source(s): Authors' own work

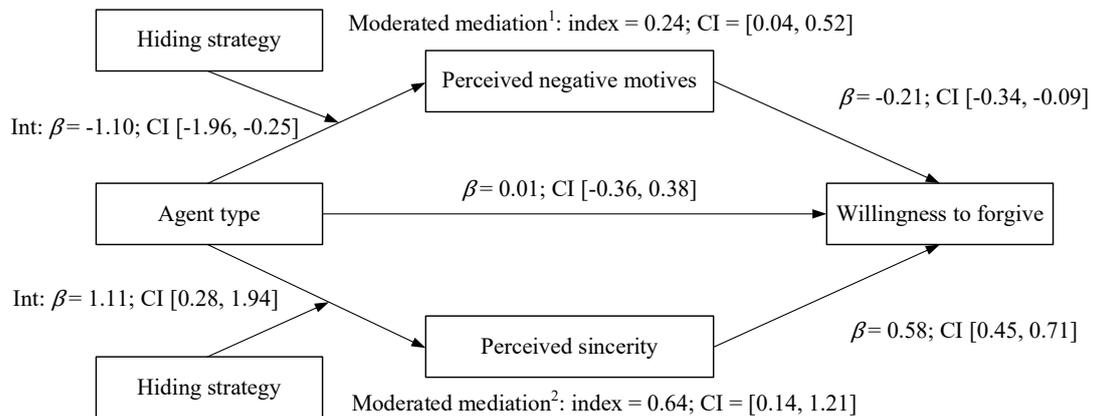
Plate 3: Study 2: Willingness to forgive as a function of agent type and hiding strategy



Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; Error bars = ± 1 SEs.

Source(s): Authors' own work

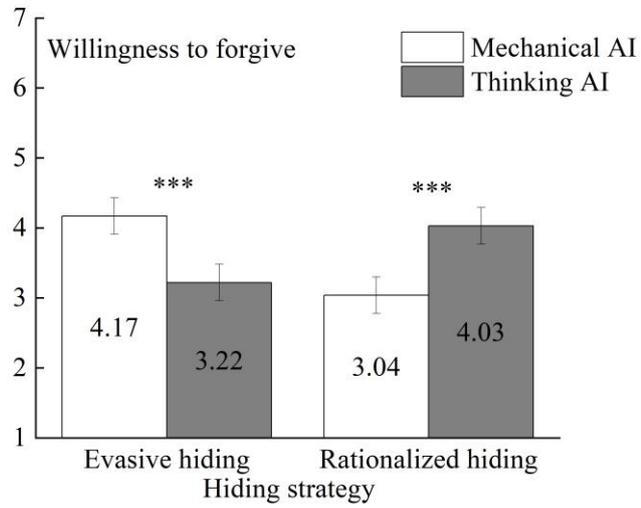
Plate 4: Moderated mediation analysis



Int: Interaction term (Agent type \times Hiding strategy)
 Moderated mediation¹: Perceived negative motives as a mediator
 Moderated mediation²: Perceived sincerity as a mediator

Source(s): Authors' own work

Plate 5: Study 3: Willingness to forgive as a function of agent type and hiding strategy



Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; Error bars = ± 1 SEs.

Source(s): Authors' own work

Table 1. Distinguishing dimensions of evasive hiding, rationalized hiding, and related service recovery strategies

Strategy category	Literature source	Concealment of true intention	Future action commitment	Cost type (economic / symbolic)	Primary user perception triggered
Evasive hiding	Connelly et al. (2012); Gong et al. (2025)	High (strong intentional concealment)	Yes (ambiguous or false)	None	Uncertainty; misleading expectations
Rationalized hiding	Connelly et al. (2012); Gong et al. (2025)	Moderate (partial concealment)	No	None	Perceived reasonableness; uncontrollable attribution
Apology	Javornik et al. (2020)	No	Yes (genuine)	Low (if without compensation)	Satisfaction; justice
Compensation	Tang et al. (2018)	No	No (action already undertaken)	High (economic cost)	Satisfaction; perceived goodwill
Low-competence apology	Antonetti and Baghi (2023)	No	Yes	Symbolic cost	Perceived apology costliness
Explanatory response	Zhao et al. (2020)	No	Sometimes	None	Transparency; causal understanding
Affiliative humor	Béal and Grégoire (2022)	No	No	None	Warmth; appreciation
Aggressive humor	Béal and Grégoire (2022)	No (but induces hostility)	No	None	Hostility; disrespect

Source(s): Authors' own work

Table 2. Results of moderated mediation effects (Study 2)

Model	Perceived negative motives (PNM)			Perceived sincerity (PS)			Willingness to forgive (WTF)		
	β	t	95%CI	β	t	95%CI	β	t	95%CI
Constant	1.46	3.01**	(0.51, 2.42)	4.11	8.70***	(3.18, 5.04)	3.04	7.59***	(2.25, 3.83)
Agent type	1.74	2.53*	(0.38, 3.09)	-1.56	-2.33*	(-2.87, -0.24)	0.01	0.05 ^{ns}	(-0.36, 0.38)
Hiding strategy	1.23	4.03***	(0.63, 1.83)	-0.09	-0.29 ^{ns}	(-0.67, 0.50)	-	-	-
Interaction term	-1.10	-2.56*	(-1.96, -0.25)	1.11	2.64**	(0.28, 1.94)	-	-	-
Perceived negative motives	-	-	-	-	-	-	-0.21	-3.48***	(-0.34, -0.09)
Perceived sincerity	-	-	-	-	-	-	0.58	9.03***	(0.45, 0.71)
Indirect effect ¹ (Agent type→PNM→WTF)				Effect		SE			95%CI
Evasive hiding				-0.13		0.08			(-0.34, -0.01)
Rationalized hiding				0.10		0.07			(-0.03, 0.27)
Indirect effect ² (Agent type→PS→WTF)				Effect		SE			95%CI
Evasive hiding				-0.26		0.17			(-0.62, 0.07)
Rationalized hiding				0.38		0.19			(0.03, 0.77)

Notes: ns = not significant; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Source(s): Authors' own work

Appendix

Appendix A: Stimuli

Studies 1, 2

AI condition



This is your exclusive AI assistant, named Xiao Hai. He will handle your inquiries. Subsequently, you'll see how he responds to your feedback.

Source(s): Authors' own work

Human condition



This is your exclusive assistant, named Xiao Hai. He will handle your inquiries. Subsequently, you'll see how he responds to your feedback.

Study 3

Mechanical AI



This is your exclusive agent, named Xiao Hai, based on a basic artificial intelligence technology natural language processing model, capable of handling most basic conversation scenarios. We hope it can satisfy your needs.

Source(s): Authors' own work

Thinking AI



This is your exclusive agent, named Xiao Hai, based on a deep learning natural language processing model, trained to generate human-like conversations and understand the context of conversations to generate coherent responses. We hope it can satisfy your needs.

Appendix B: Measurement items

Variables	Measurement Items	Sources
Evasive hiding Studies 1, 2, 3	The agent promised to help, but I believe they never really intended to do so.	Connelly et al. (2012)
	The agent promised to help later, but I think they are just stalling.	
	The agent provided some information, but it's not what I really wanted.	
Rationalized hiding Studies 1, 2, 3	The agent hinted that they could not help.	Connelly et al. (2012)
	The agent suggested that this issue could not be resolved due to certain reasons.	
	The agent implied that this issue could not be resolved due to reasons beyond their control.	
Willingness to forgive Studies 1, 2, 3	Given [Agent name]'s response, I will condemn the company.	Wei and Ran (2019)
	Given [Agent name]'s response, I will forgive the company.	
	Given [Agent name]'s response, I let go of the resentment toward the company.	
	Given [Agent name]'s response, I let go of the negative feelings I had toward the company.	
Perceived negative motives Studies 2, 3	This agent has good intentions (1) - has bad intentions (7).	Joireman et al. (2013)
	This agent (AI/human) does not intend to exploit users (1) - intends to exploit users (7).	

Variables	Measurement Items	Sources
	This agent (AI/human) is primarily driven by user interests (1) - is primarily driven by self-interest (7).	
Perceived sincerity Studies 2, 3	This agent is sincere. This agent is emotional. This agent seems to be concerned about the issue.	Barasch et al. (2016)
Clarity Study 2	This agent's response is clear. This agent's response is unclear.	Antonetti and Baghi (2023)
Credibility Study 2	This agent's response is realistic. This agent's response is unrealistic.	Antonetti and Baghi (2023)
Benevolent intentions Study 2	This agent had its own benevolent intentions. This agent had its own generous intentions. This agent had its own good intentions.	Garvey et al. (2023)
Selfish intentions Study 2	This agent had its own selfish intentions. This agent had its own greedy intentions. This agent had its own bad intentions.	Garvey et al. (2023)
AI capabilities	This agent is similar to a human.	Garvey et al.

Variables	Measurement Items	Sources
Study 3	This agent is able to think like a human.	(2023)
	This agent is able to behave like a human.	
Social presence	I felt a sense of human contact with the agent.	Verhagen et al.
		(2014)
Study 3	I felt a sense of personalness with the agent.	
	I felt a sense of human sensitivity with the agent.	

Appendix C: The reliability and validity assessment results

Drawing on the approach adopted by Huang et al. (2025), we examined the reliability and validity of the measurement items across the studies in this research. Internal consistency was evaluated using Cronbach's alpha and composite reliability (CR). While Cronbach's alpha represents a relatively conservative reliability estimate, composite reliability accounts for variations in item loadings. Following established guidelines, values exceeding 0.70 indicate acceptable reliability (Hair et al., 2016). As reported in Table 1, both Cronbach's alpha and CR for all constructs surpass this recommended threshold.

Convergent validity was examined through average variance extracted (AVE) and standardized factor loadings. AVE values greater than 0.50 suggest adequate convergent validity (Hair et al., 2016), a condition satisfied by all constructs in the proposed model (see Table 1). Therefore, the results provide evidence supporting the convergent validity of the measurement model.

Lastly, we calculated the heterotrait–monotrait ratio (HTMT) for the three studies. Values exceeding 0.90 indicate a lack of discriminant validity (Hair et al., 2016), while a more conservative threshold of 0.85 has also been suggested (Henseler et al., 2015). As shown in Table 2 (Study 1), Table 3 (Study 2), and Table 4 (Study 3), all HTMT values across the studies are well below these thresholds, providing evidence for the discriminant validity of the constructs.

Table 1. Cronbach's alpha, composite reliability (CR), average variance extracted (AVE)

Study	Construct	Cronbach's alpha	CR	AVE
Study 1	Evasive hiding	0.883	0.889	0.731
Study 1	Rationalized hiding	0.874	0.878	0.705
Study 1	Willingness to forgive	0.888	0.893	0.680
Study 2	Evasive hiding	0.779	0.809	0.601
Study 2	Rationalized hiding	0.863	0.862	0.678
Study 2	Willingness to forgive	0.951	0.953	0.836
Study 2	Perceived negative motives	0.882	0.887	0.723
Study 2	Perceived sincerity	0.887	0.893	0.736
Study 2	Clarity	0.860	0.867	0.765
Study 2	Credibility	0.844	0.851	0.740
Study 2	Benevolent intentions	0.930	0.933	0.823
Study 2	Selfish intentions	0.962	0.962	0.894
Study 3	Evasive hiding	0.912	0.961	0.784

Study	Construct	Cronbach's alpha	CR	AVE
Study 3	Rationalized hiding	0.889	0.892	0.733
Study 3	Willingness to forgive	0.949	0.950	0.828
Study 3	Perceived negative motives	0.884	0.883	0.718
Study 3	Perceived sincerity	0.940	0.941	0.843
Study 3	AI capabilities	0.817	0.820	0.604
Study 3	Social presence	0.952	0.954	0.873

Source(s): Authors' own work

Table 2. Heterotrait-monotrait ratio (Study 1)

	Evasive hiding	Rationalized hiding	Willingness to forgive
Evasive hiding	-		
Rationalized hiding	0.108	-	
Willingness to forgive	0.282	0.342	-

Source(s): Authors' own work

Table 3. Heterotrait-monotrait ratio (Study 2)

	Evasive hiding	Rationalized hiding	Willingness to forgive	Perceived negative motives	Perceived sincerity	Clarity	Credibility	Benevolent intentions	Selfish intentions
Evasive hiding	-								
Rationalized hiding	0.212	-							
Willingness to forgive	0.384	0.417	-						
Perceived negative motives	0.125	0.121	0.415	-					
Perceived sincerity	0.520	0.081	0.643	0.386	-				
Clarity	0.331	0.103	0.285	0.183	0.393	-			
Credibility	0.306	0.126	0.260	0.227	0.333	0.721	-		
Benevolent intentions	0.474	0.324	0.742	0.484	0.637	0.429	0.380	-	
Selfish intentions	0.498	0.352	0.664	0.443	0.559	0.312	0.217	0.905	-

Source(s): Authors' own work

Table 4. Heterotrait-monotrait ratio (Study 3)

	Evasive hiding	Rationalized hiding	Willingness to forgive	Perceived negative motives	Perceived sincerity	AI capabilities	Social presence
Evasive hiding	-						
Rationalized hiding	0.317	-					
Willingness to forgive	0.576	0.423	-				
Perceived negative motives	0.604	0.507	0.732	-			
Perceived sincerity	0.659	0.269	0.740	0.745	-		
AI capability	0.195	0.094	0.342	0.271	0.396	-	
Social presence	0.606	0.321	0.761	0.731	0.836	0.470	-

Source(s): Authors' own work