Print ISSN: 1738-3110 / Online ISSN 2093-7717
JDS website: http://www.jds.or.kr/
http://dx.doi.org/10.15722/jds.22.02.202402.103

Consumers' Tolerance When Confronted with Different Service Types in Service Retailing*

Chengcheng YU[1], Na CAI[2], Jinzhe YAN[3], Yening ZHOU[4]

Received: November 03, 2023. Revised: December 14, 2023. Accepted: February 05, 2024.

Abstract

Purpose: With the popularity of artificial intelligence (AI) in the service industry and occurrence of service failures in AI-based services, understanding human-robot interaction issues in service failure situations is especially important. Some issues which deserve further empirical investigation are whether consumers can develop the same tolerance for chatbots after service failure as they have for human agents, and the relationship between agent type and tolerance is mediated by the mechanisms of perceived warmth and perceived competence. **Research Design, Data, and Methodology:** This research experimentally collected and analyzed data from 119 university students who had experienced chatbots service failures. Differences in tolerance towards human agents and chatbots after experiencing service failures were explored, with a further examination of the mediating pathways between this relationship via perceived warmth and perceived competence. **Results:** Consumers are more tolerant of service failure with chatbots compared to service failure with human agents. Significant mediation of the relationship between service agent and service failure tolerance by perceived competence, while perceived warmth has no significant mediating effect. **Conclusions:** This research enhances our understanding of AI-assisted services, human-computer interaction, improves the service functionality of existing smart devices, and deepens the understanding of the relationship between consumer responses and behaviors.

Keywords: Service failure, Chatbots in marketing channels, Perceived warmth, Perceived competence

JEL Classification Code: M31, L81, O33

1. Introduction

Artificial Intelligence (AI), technological innovations, and a wide range of messaging platforms have contributed to the large-scale adoption of chatbots by service providers as primary service agents and enhanced the level of human-robot interaction patterns with consumers (Følstad et al., 2018). With AI and machine learning technology advancements, computer program-based chatbots are now learning to mimic human conversations through multiple modes of communication, such as voice and text, to

* This work was supported by the Gachon University research fund of 2023 (GCU-202303790001).
1 First Author. Ph.D. Student, Department of Business Administration, Gachon University, South Korea. Email: 202255041@gachon.ac.kr
2 Second Author. Ph.D. Student, Department of Business Administration, Gachon University, South Korea. Email: 202155074@gachon.ac.kr
3 Third Author & Corresponding Author. Assistant Professor, College of Business, Gachon University, South Korea. Email: yanjz@gachon.ac.kr
4. Fourth Author & Corresponding Author. Ph.D. Student, Department of Business Administration, Gachon University, South Korea. Email: Yeningzhou12@gmail.com

successfully provide services to consumers (Luo et al., 2019). Chatbots allow businesses to reduce labor costs, enhance efficiency, and provide personalized and customized service capabilities. Crucially, chatbots continue to revolutionize and optimize the tools and techniques of human-computer chatbots continue to revolutionize and optimize the tools and techniques of human-computer interactivity (Lv et al., 2021; Lv et al., 2022). Companies have enhanced human-robot engagement in online service environments to improve the consumer experience. Hence, in the modern business environment, chatbots are increasingly being deployed in online service processes and used as substitutes for human service agents.

As a technology, chatbots, unlike human agents, are not without flaws (Choi et al., 2021). As they do not perfectly reproduce the way people communicate fluidly with each other, chatbots often fail to meet consumer expectations (Sheehan et al., 2020). Under critical situations, this may negatively impact the consumer's service experience and lead to the service's failure to deliver the desired results (hereafter, service failure) (Crolic et al., 2022). A lack of successful experiences with various chatbot-based services may adversely affect consumer attitudes towards interaction and subsequent actions for such services. For example, the Henn-Na Hotel in Japan said it was the first hotel of its kind with robots, dismissed over 100 robot machines. The reason was that they were producing more problems than they were solving, with the hotel's intelligent robots interpreting guests' snores as signals for assistance and providing wake-up calls several times during the night (Lv et al., 2021).

Most studies explore how robotic service failure affects consumer attitudes from the perspectives of both attribution and remediation strategies. Few examine the variability in consumer psychological evaluations of different types of service agent failures. When there is a problem with a service, consumers may judge whether the service has gone wrong based on their feelings, which can further impact their subsequent behaviors and attitudes. Chatbots have become a topic of interest as a new type of service provider. Therefore, understanding consumers' tolerance toward different types of service agents (i.e., humans versus chatbots) after suffering service failure is important. This can help businesses engage in precise remediation measures in a more efficient and cost-effective manner, which in turn facilitates collaboration between consumers and businesses to improve the customer remediation.

Based on attribution theory, we argue that when consumers experience service failures while using a chatbot, they are more inclined to blame the company for providing the service for the failures and demonstrate a relatively high level of tolerance for the chatbots. In contrast, when a human agent's service fails, consumers tend to mistakenly believe that it is caused by the human agent's behavior. Consequently, consumers have relatively low tolerance for such human agent-induced service failures.

Furthermore, based on the Stereotype Content Model (SCM), we present a view of the universal perceptual dimensions of social individuals: perceived warmth and perceived competence. This provides a useful theoretical framework for understanding consumer tolerance towards humans and chatbots service agents in the event of service failures. People's judgment of others is affected by two fundamental factors: warmth and competence. Warmth refers to the traits associated with perceived intentions, including friendliness, helpfulness, sincerity, trustworthiness, and morality. Competence refers to the traits associated with perceived abilities, including intelligence, skills, creativity, and efficacy (Fiske et al., 2007). While these two concepts are frequently researched in the context of people's social judgments, they are also associated with the judgments of other entities, such as robots and virtual agents (Aaker et al., 2010; Bergmann et al., 2012). When they interact with chatbots, consumers develop different perceptions of warmth or competence for the chatbots. Despite the growing number of potential applications of chatbots, few empirical works investigate how consumer attitudes are affected when they have different perceptions of chatbots. Here, we examine whether the dimensions of consumers' perceptions of a service agent after a service failure may affect their tolerance.

In summary, building on attribution theory and the SCM, this study explores how consumer perceptions affect consumer attitudes towards a service agent (humans versus chatbots) in the context of service failure. Specifically, we focus on: (1) comparing consumer tolerance towards different service agents (humans versus chatbots) after service failure; and (2) perceived warmth and perceived competence act as mediators in the connection between type of service agents (humans versus chatbots) and consumer tolerance after service failure. We conducted an experiment to investigate these topics. The results suggest that perceptions of warmth and competence explain the impact of the type of service agents on tolerance.

This study makes several contributions to existing literature. First, this study expands on AI service studies from exploring consumer acceptability and receptivity to AI services to exploring human-robot interaction issues in service failure contexts. This not only enriches the content of customer experience, servant and dialogue marketing but also expands the scope of research in the field of chatbots. Second, we test the effects of two types of service agents—humans and chatbots agents—on consumer tolerance. Thus, we extend research on human-computer interaction and innovatively apply related theories to human-computer interaction experiments. Finally, based on the SCM, this study reveals the mediation effects of perceptions of perceived warmth and perceptions of perceived competence on consumer tolerance. Thus, we provide a theoretical foundation for enhancing the service agent attributes to provide better recovery after service failure. Furthermore, we extend the application of the SCM theory to the context of AI service failure.

2. Literature Review and Hypothesis Development

2.1. Chatbots in Marketing Channels

Chatbots are computer programs that can be used to satisfy users by engaging in text-based conversations with them (Wang & Zhang, 2018). Studies demonstrate that AI systems make more precise and higher performing decisions than traditional human-operated ones (Batra & Antony, 2001). As chatbots can mimic the tone and behavioral patterns of humans in conversations, consumers tend to instinctively assume that these chatbots may very well perform several tasks similar to human agents (Hong & Williams, 2019). Consequently, chatbots are endowed with the unique attribute of "human dehumanization and technological humanization" (Kaczorowska-Spychalska, 2019). With advances in the capabilities of AI and natural language processing systems, chatbots are being used as chat platforms for personal, professional, and business communication (Følstad & Brandtzæg, 2017). Essentially, they have become an example of one of the most widespread applications of modern AI. Chatbots have been applied in various application domains, with customer service being the most common. Chatbots can be configured as a separate support gateway or as frontline support; however, when a chatbot encounters a challenge that it cannot handle on its own, it is authorized to contact a skilled human service agent (Wilson & Daugherty, 2018).

While studies have extensively researched chatbots design and consumers' experiences with them, these studies are based on the assumption of positive functional interactions between consumers and chatbots. These assumptions can be explained psychologically, and to some extent, reflect the gap between what people want from chatbot functionality and what is actually happening. As the application domain of chatbots expands, the problems that accompany them also expand, the most significant of which is the provision of unsuccessful services. Consequently, a new wave of research has recently begun analyzing chatbots service failures, which occur mainly when chatbots do not meet consumer expectations in terms of service standards. In reality, it is not uncommon for chatbots to fail to understand consumers' input, and thus, fail to live up to their expectations. Moreover, consumers may very well become reluctant to re-adopt a chatbot service later if the chatbot has incorrectly responded before (Sheehan et al., 2020). This is the largest obstacle encountered in the use of chatbots. Surprisingly, while chatbots increasingly replace human workers in a broad array of service sectors and human-robot interactions become more frequent, little is understood regarding the experience of users after a service failure, making it particularly vital for understanding how consumers react to chatbots.

2.2. Service Failure and Consumer Tolerance

Service failure occurs when service delivery is below than customer expectations (Newton et al., 2018). Service failures often trigger negative consumer emotions, such as dissatisfaction, complaints, anger, and even retaliatory behaviors (Tsarenko et al., 2019). As the scope of tasks that AI can take on expands, chatbots may also face various kinds of service failures. From the point of view of service needs, service failures are categorized into two types: core and interaction service failures (Keaveney, 1995). When a core service is not in line with the fundamental requirements of the consumer, a serious mistake or flaw exists in the service. Unsuccessful interactive services are related to the attitudes and actions of workers in their interactions with customers, which are problems in the process of service provision, such as cold treatment or long wait times (Sparks & Browning, 2010).

Meanwhile, two categories of service failures have been identified from a service phase point of view: outcome and process failures (Smith et al., 1999). An outcome failure occurs when a service fails to satisfy the fundamental requirements of consumers. A failure in the process occurs when there are flaws or deficiencies in the delivery of the service. In a process failure, human-like robots do not live up to customer demands, which may reduce customer satisfaction. In contrast, when confronted with a failed outcome, consumers have comparable expectations for both human-like and non-human-like robots and do not exhibit significant disparities in satisfaction levels with non-human-like robots (Choi et al., 2021). Although chatbot applications are rapidly expanding, their functionality is not flawless; consequently, service failures are unavoidable. Moreover, expecting chatbots to demonstrate excellence in all areas seems unrealistic. Service failures can be a source of substantial damage to the service provider's business, such as loss of customers (Sun, 2021) and unfavorable information circulated through oral communication (Crisafulli & Singh, 2017). Despite the gradual increase in chatbots usage in the service sector, research on consumer reactions to and attitudes towards chatbots after failed service delivery remain scarce. We need further investigations on whether users want to hold bot service agents accountable just as they would traditional humans.

Attribution theory was formulated by social psychologists (Heider, 2013). It states that people's attempts to understand the cause of specific events through their own efforts to see why these events occur. When individuals perceive bad consequences, they will actively search for the reasons for these consequences and for the people who can be held accountable for them (Choi & Mattila, 2008). The theory of attribution is extensively used in research on how customers respond to failure, including interactions between humans and robots, as well as chatbots. Building on this theory, we claim the perceived control of chatbots service agents, as opposed to that of human service agents, is crucial in determining the manner in which consumers assign responsibility for service failures.

Chatbots operate based on computer algorithms programmed by humans to perform their operations. Further, the bot service provider company has less degree of control over the result compared to using a human service agent (Hong & Williams, 2019). Compared to humans, computational systems are deficient in emotional processing, initiative, and independent thinking (Gray et al., 2012). Precisely because robots lack a clear subjective intention, they are not able to perform with purpose, which can lead to service failures; therefore, the responsibility for failures cannot simply be attributed to them. In such a context, consumers are more likely to proactively seek external participants to take responsibility for the consequences of failure (Weiner, 2000). When AI Services fail to work, the service delivery businesses are blamed for providing a failure of service (Leo & Huh, 2020). In the majority of cases, chatbots are considered highly compliant. This is due to the control, or a lack of it, they have been endowed with by the service provider and its developers, and the relatively small amount of responsibility chatbots have to bear compared to a human service agent. When a chatbot, rather than a human agent, causes service failure, people are more likely to blame the service provider as they perceive that the provider has control over the chatbots. Similarly, when AI-based services experience failures, consumers will be more inclined to target the service provider rather than the chatbots because the chatbots would never be held liable for its failures (Gray & Wegner, 2012). In contrast, if a human service representative is involved in a service failure, consumers may place responsibility on the human agent perceived to be directly responsible for the failure.

Based on this literature and current theories, we argue that, compared to human agents, where service companies may be held accountable for service failures, chatbots are largely perceived to exhibit weaker control over themselves and may be blamed less for service failures. Therefore, after a consumer has a failed service experience with a chatbot, the service supplier is blamed for the failure and show a higher tolerance for chatbots. In contrast, after a service failure, customers tend to attribute the cause of such failure to the human agent's lack of competence or negligence (Belanche et al., 2020). Hence, consumers have a lower tolerance for service failures by human agents. Accordingly, we propose the following hypothesis:

H1: Consumers are more tolerant of service failures caused by chatbots agents than by human agents.

2.3. Perceived Warmth and Perceived Competence

On the basis of the SCM, warmth and competence are defined as two of the core aspects of human perception. These two dimensions form the underlying social structural framework for helping to understand the ways in which people sense and react to issues in society. For instance, using these two dimensions, people can describe others, and quickly categorize them as friendly or threatening (Fiske et al., 2007). Since these views are deeply embedded in the evolutionary history of humans, social evaluations of warmth and competence have been prevalent across a broad span of interactions in the social world. Further, these dimensions are capable of elucidating major differences in views of everyday social behavior. Scholars generally recognize that these dimensions emerged partly because of the evolutionary adaptability of such perceptions, allowing people to rapidly determine others' intentions, trustworthiness, and potential threats. Essentially, the framework for warmth and competence is a theoretical system which focuses on human relationships and team effectiveness. Here, warmth is viewed as "perceived intentions" and competence as "perceived abilities." When a person demonstrates greater competence and warmth, others perceive them more favorably (Wojciszke et al., 2009; Wortman & Wood, 2011). Specifically, when people are seen as warm or competent, they are more likely to elicit positive feelings and actions from others, while people who are seen as less warm or competent are more likely to elicit negative responses (Fiske et al., 2007). Several psychological studies indicate that warmth and competence do help mold people's responses in various environments and in pursuit of various goals (Cuddy et al., 2008).

The processes of social cognition extend beyond human relationships and can also manifest during interaction with nonhuman entities displaying human-like traits (Mieczkowski et al., 2019; Oliveira et al., 2019). Indeed, scholars and administrators have gradually increased their focus on perceptual dimensions, which has contributed to a more in-depth understanding of how consumers evaluate and respond to service providers such as chatbots. In contexts where chatbots are used extensively, warmth relates to the robot's perception of human kindness or malice, whereas competence concerns whether the robot has the technology and ability to perform the agent's tasks. Studies have discussed the perceptions of chatbots' warmth and competence. Research shows that humans react to chatbots in the same manner as people react to each other (Reeves & Nass, 1996). Robots were initially created to accomplish utilitarian goals; therefore, competence was considered an intrinsically associated trait. Meanwhile, as robots evolve, they will be humanized with more and more humanlike characteristics, they exhibit a feeling of warmness (Kim et al., 2019). The latest type of human-machine interaction not only triggers the social response that often occurs in human-to-human interactions, but also perceptual mechanisms that enable people to perceive the robots' presence with respect to their warmth and competence (Belanche et al., 2021). Chatbots utilize physical characteristics and verbal gestures resembling those of humans so that consumers can feel warmth during their interactions with chatbots. Furthermore, as chatbots were originally designed to serve as utilities and provide more efficient services to consumers, competence is considered as an attribute that should be present and is expected in chatbots. Consumers' perceptions of competence are unrelated to the degree of humanization of the bot itself (Choi et al., 2021; Kim et al., 2019). Multiple studies have demonstrated that, when people evaluate chatbots, they typically use the of warmth and competence dimensions; these perceptual factors further influence their emotional responses to chatbots (Mieczkowski et al., 2019). Therefore, this study introduces two mediating mechanisms–perceived warmth and perceived competence–into the conceptual model.

Perceived warmth refers to the perception of other people's intentions. It captures traits that contribute to maintaining relations and social functioning, such as kindness, affability, cooperativeness, genuineness, conviviality, joy, and humor (Aaker et al., 2010; Cuddy et al., 2008; Kirmani et al., 2017). When the social presence of AI-based entities is strong, consumers often mistakenly assume that they are communicating with real people rather than using virtual AI entities (Lee et al., 2006). This decreases the social distance between the consumers and chatbots, thereby strengthening their bond. When consumers perceive warmth from chatbots, they are more actively engaged, and willing to interact and collaborate with the bot in-depth. Perceived warmth not only helps foster positive perceptions of chatbots, but also provides a conducive atmosphere for collective value creation. Consumers who feel the warmth are more likely to trust the chatbot's competence, develop a deeper emotional bond with them, and perceive that they are instinctively attentive to their perceived requirements (Babel et al., 2021). Warmth perception can convince consumers that they will receive positive feedback from the chatbots for their shared needs. Research confirms that perceived warmth is associated with happiness and positive behavioral expectations after bot failure (Choi et al., 2021). Further, perceived warmth towards AI assistants is positively correlated with consumer tolerance (Lv et al., 2021). In a service failure scenario, consumers may develop greater tolerance if they perceive warmth from the service agent. Accordingly, we make the hypothesis that:

H2a: Perceived warmth mediates the relationship between service agents (humans versus chatbots) and service failure tolerance (consumers' willingness to continue using the service after a service failure).

Perceived competence reveals various perceptual abilities. It encompasses traits pertinent to task performance, including effectiveness, intelligence, proficiency, ability, skill, and task accuracy (Cuddy et al., 2008; Kirmani et al., 2017). In the AI context, developers are working on incorporating human-like characteristics into AI devices with the intention of allowing chatbots to exhibit superior intelligence, making their performance comparable to that of an average human employee and perhaps even more advanced (Chi et al., 2020). Human similarity tends to closely correlate with the level of proficiency and cognitive ability displayed in chatbots tasks (Lu et al., 2021). Consumers are more inclined to build relationships with chatbots because they perceive them to be efficient and effective in information search, or reliable and trustworthy (Rajaobelina et al., 2021). Chatbots demonstrate a range of behaviors that may enhance consumers' awareness and perception of the chatbots' functionality. During the interaction, various behaviors of the chatbots, such as providing key information, giving valuable suggestions, and fulfilling the user's needs as much as possible (Liu et al., 2018), can be seen as symbolizing their efforts to accomplish things and better deliver high-quality service. Assuming that chatbots can proactively meet consumers' needs and demonstrate a high degree of proficiency during interactions, consumers are likely to perceive these bots as entities with appropriate competence. In addition, chatbots behavior demonstrates a positive attitude in that it can predict and positively respond to customer needs and requests. This conveys a feeling of self-confidence and effectiveness that leads participants to believe that the chatbots are capable of fully understanding and maximizing the satisfaction of consumer expectations.

Consumers' perceptions of a chatbot's competence are further enhanced when they notice that the chatbots responds affirmatively and accurately, which further strengthens their confidence in the chatbots' proficiency and expertise. During service failures, consumers pay close attention to and expect the service agent to resolve the issues efficiently and effectively. Simultaneously, consumers can analyze the intrinsic competence of a service agent to predict their tolerance and satisfaction levels towards service failures (Liu & Li, 2022). Consumers will choose to continue using the service if they feel that the service agent is capable of successfully restoring it (Lv et al., 2022). Consumers may be more tolerant of a service failure if the service agent allows them to perceive its competence. Accordingly, we make the hypothesis that:

H2b: Perceived competence mediates the relationship between service agents (humans versus chatbots) and service failure tolerance.

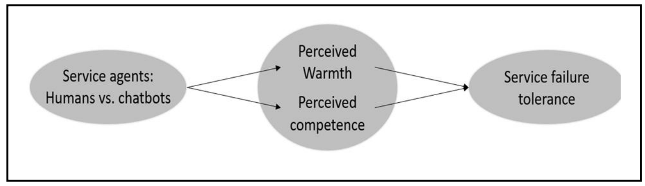
3. Research Method

3.1. Design of Experiment

We conducted a service agent (human versus chatbots) experiment to test the model (see Figure 1). Differences in consumer tolerance between the two types of service agents were assessed in a service failure scenario in which the service agent (humans versus chatbots) failed to successfully solve the consumer's problem. Further, we evaluated the mediating function of perception of warmth and perception of competence. To improve external validity, 119 university students who had experienced chatbot service failures were selected as participants using a real-time chatbots via Chatplat, a professional online chatbot platform.

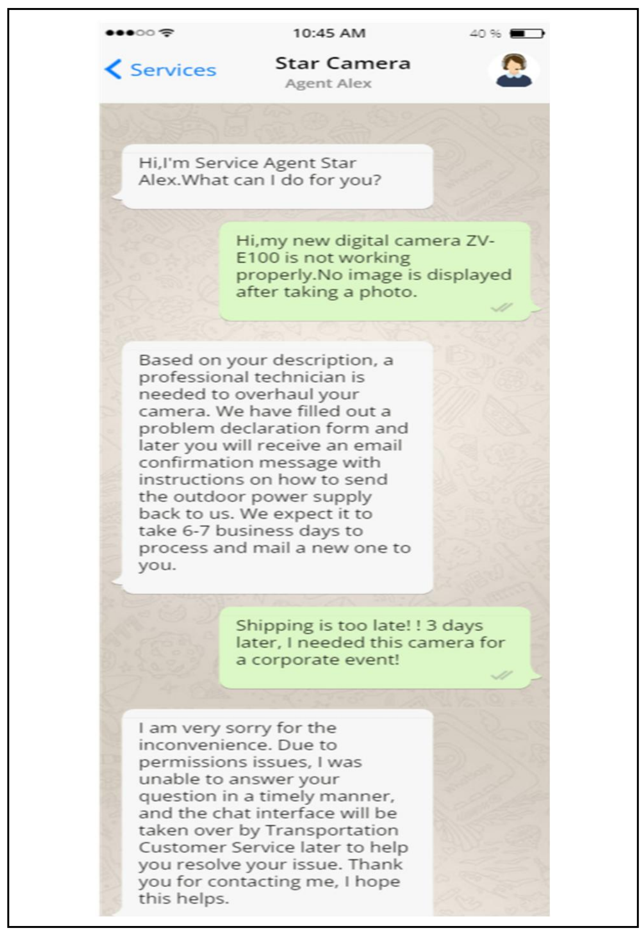
Adapting previous literature (Crolic et al., 2022) and the actual phenomena often found in service failures, a scenario was designed (see Figure 2 & Figure 3). Participants were asked to watch a scenario in which they negotiated with an online customer service for a replacement digital camera and to imagine a scenario in which they interacted with a service agent. The scenario was as follows: participants purchased a digital camera for an upcoming group activity, and when consumers received the camera and tried it out, they found that the camera was not working properly and then called customer service, where they were transferred to multiple customer service representatives and told that they would have to wait in line to be accommodated. When the consumer finally engaged in a negotiated settlement conversation with customer service regarding a replacement digital camera, it was learned that a new power supply would take about seven days to replace and was expected to arrive after the start of the group activity, which went against the participant's plan. When they attempted to contact customer service by phone again, customer service responded that they were unable to help and suggested contacting the online customer service center.

After reading the script, participants were asked to answer questions related to the realism of the situation and were then randomly assigned to a real-time dialog box with an online customer service agent (either human customer service agent or chatbot) to address the issue of replacing the outdoor power supply. Subsequently, participants were asked to complete a measurement questionnaire. Finally, participants answered demographically relevant questions.



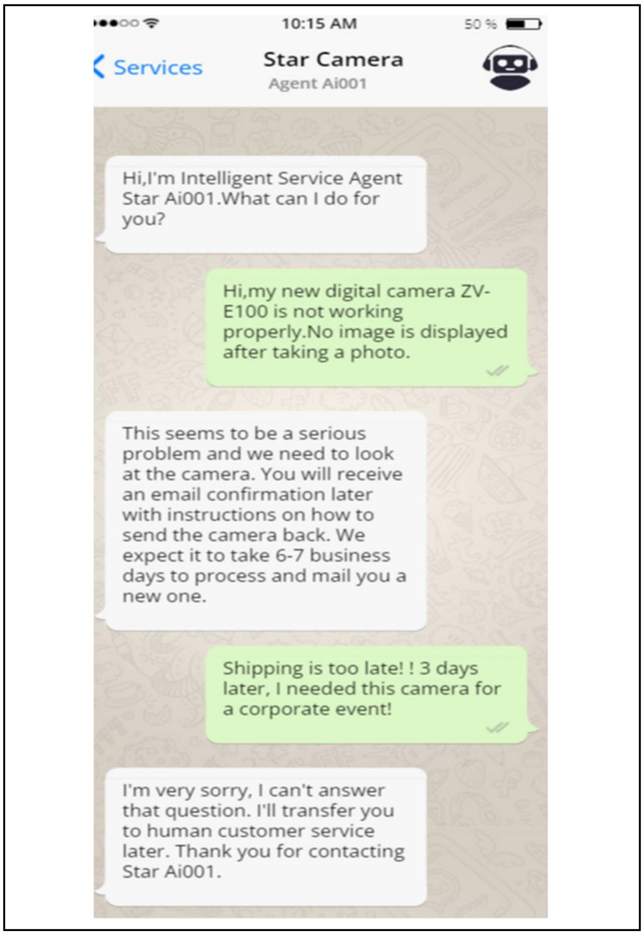
Source: Authors generated the above Figure

Figure 1: Research Model



Source: Authors generated Figure 2 via www.fakewhats.com to illustrate the procedure of the experiment

Figure 2: Human Customer Service Agent



Source: Authors generated Figure 2 via www.fakewhats.com to illustrate the procedure of the experiment

Figure 3: Chatbot Customer Service Agent

3.2. Operational Inspection

We manipulated service failure scenarios, scenario realism, and agent type. Service failure scenario was measured using four items: "When confronted with a service failure situation it makes me feel: sad/happy; bad/mood-good/mood; irritable/pleased; depressed/cheerful" (Townsend & Sood, 2012). The ANOVA results indicated that all participants exhibited high levels of anger in response to the situational experience ($M_{human} = 4.023$ versus $M_{chatbot} = 4.223$, $F(1,117) = 2.247$, $p = 0.137 > 0.05$), with no significant difference between the groups. Therefore, the manipulation of the context was successful.

Scenario realism was measured by using the following three items: "Please rate this scenario: this scenario is real; this scenario is credible; it is easily possible for me to imagine myself as the customer" (Pavone et al., 2023). The ANOVA results indicated that the participants perceived the context to be realistic ($M_{human} = 4.047$ versus $M_{chatbot} = 4.230$, $F(1, 117) = 3.322$, $p = 0.071$).

Agent-type source attribute manipulation was used to measure questionnaire validity by employing the following three items: "Do you think you are having a conversation with ___: options include 'chatbot', 'human customer service agent', and 'unsure'" (Lou et al., 2022). The ANOVA results indicated that participants randomly assigned to the human customer service perceived this customer service agent as more human-like ($M_{human} = 3.578$ versus $M_{chatbot} = 1.418$, $F(1, 117) = 84.538$, $p = 0.000$), while those randomly assigned to the robot customer service perceived this customer service agent as more robot-like ($M_{human} = 2.484$ versus $M_{chatbot} = 4.636$, $F(1, 117) = 80.656$, $p = 0.000$).

3.3. Measures

Participants followed instructions from the telephone customer service and entered a real-time dialog box with the online customer service to address the two issues experienced during the replacement of the power supply with either a human customer service or chatbot. Participants then completed the questionnaire described below, with all items evaluated on a five-point Likert scale, with "1" indicating "strongly disagree", and "5" indicating "strongly agree").

Perceived warmth was measured using the four response items to the question: "Do you think this online agent is: friendly/enthusiastic/kind/sincere". Perceived competence was measured using the following four response items: "capable/competent/intelligent/skilled" (Aaker et al., 2010; Wang et al., 2017; Xu et al., 2023).

Customer tolerance was measured using four operational check questions based on previous literature: "Please rate this experience: this service failure is inexcusable / you would communicate that failure to others / you would be prepared to recommend an intelligent customer service agent to others / you would be willing to have an intelligent customer service agent help you solve your problem again" (Augusto de Matos et al., 2009; Lv et al., 2021).

3.4. Analysis Method

Analysis of variance: ANOVA is a collection of statistical models and their associated estimation procedures (such as the "variation" among and between groups) used to analyze the differences among means. ANOVA was developed by the statistician Ronald Fisher (Fisher, 1919). ANOVA is based on the law of total variance, where the observed variance in a particular variable is partitioned into components attributable to different sources of variation. In its simplest form, ANOVA provides a statistical test of whether two or more population means are equal and therefore generalizes the t-test beyond two means. In other words, the ANOVA is used to test the difference between two or more means.

Mediation analysis: In statistics, mediation modeling attempts to identify and explain the mechanism or process underlying the observed relationship between the independent and dependent variables by introducing a third hypothetical variable, the mediator variable. Unlike the direct causal relationship between the independent and dependent variables, the mediation model paints a picture in which the independent variable influences the dependent variable by influencing the mediating (unobservable) variable. Thus, the role of the mediating variable is to clarify the nature of the relationship between the independent and dependent variables. The mediated effects (mediation) framework (referred to as the BK framework) proposed by Baron and Kenny has had a far-reaching impact on many social science researches such as social psychology and consumer behavior (Baron & Kenny, 1986).

4. Results

The main impact of the service agent on service failure tolerance was significant (M_{human} = 2.363 versus $M_{chatbot}$ = 2.896, F (1, 117) = 18.554, p = 0.000). Thus, consumers are more tolerant of service failures caused by chatbots than those by human agents. Therefore, H1 is supported. The impact of service agents on perceived competence was significant, and the perceived competence of human agents was higher than that of chatbots (M_{human} = 3.164 versus $M_{chatbot}$ = 2.636, F (1, 117) = 8.463, p = 0.004). Meanwhile, the impact of service agents on perceived warmth was not significant, and the perceived warmth of human agents on perceived warmth was higher than that of chatbots (M_{human} = 3.438 versus $M_{chatbot}$ = 3.123, F (1, 117) = 3.430, p = 0.067).

Mediation analysis was performed to examine the mediation impacts of perceptions of warmth and perceptions of competence using PROCESS (Model4, 5000 samples) (Hayes, 2012). The results (see Table 1) show that the mediation impact of perceptions of competence was significant (effect = -0.239, 95%BootCI = [-0.394, -0.077], not including 0), while the mediation impact of perceptions of warmth was not significant (effect = -0.146, 95%BootCI = [-0.286, 0.015], including 0). Therefore, H2b is supported, and H2a is not supported.

Table 1: Mediation Analysis Results

	Effect	SE	95%BootCI
Total	0.532	0.124	0.288, 0.777
Indirect			
Competence	-0.146	0.076	-0.286, 0.015
Warmth	-0.239	0.081	-0.394, -0.077

5. Discussion

5.1. General Discussion

With AI's emergence, the traditional way of human interaction has changed, causing significant consumer concern about it (Tussyadiah, 2020). Research on AI services has focused on customer receptivity to AI and intention to use it (Chi et al., 2022). However, AI service failures mean that customer reservations about the technology's applications require further exploration. Moreover, despite rapid advancements in AI technology, service failures remain unavoidable within the realm of the high-frequency touch industry (Lv et al., 2021; Lv et al., 2022). This study investigates consumers' responses to tolerance after experiencing perception of warmth and perception of competence, particularly in connection with service failures. Our work not only contributes a new theoretical foundation for AI service research but also expands the knowledge of service failure from human interaction to interactions between humans and AI.

Our key findings are as follows: First, in situations where a social interaction service experiences failure, consumers commonly believe that both chatbots and humans should assume some responsibility for the failure; however, chatbots assume less responsibility than humans. Because service robots are controlled through computer algorithms programmed by humans, a service provider cannot have the same level of influence over the service result as a human service provider (Hong & Williams, 2019). Research shows that, in service failures, as service robots are sometimes thought to lack control, service providers are often required to take more responsibility because they actually have control over the service robots (Gailey, 2013). Therefore, consumers exhibit higher tolerance for chatbots than human agents when confronted with the same service failure.

Second, consumer perceptions of the use of various service agents vary considerably. Specifically, human agents are generally perceived as more competent and warmer than chatbots. Consumers tend to subconsciously believe that human agents are superior to bots in problem-solving. Thus, through their chats, they generally perceive human agents to be more competent than bots. While the results for perceived warmth are not significant, consumers still experience more warmth from human agents than from chatbots.

Finally, perceived competence mediates the relationship between service agent types (human agents versus chatbots) on consumer tolerance. Most people are more inclined to believe that AI entities with human-like capabilities can perform tasks in an intelligent and competent manner (Lu et al., 2021). However, chatbots still suffer from a lack of competence compared with human agents, which leads consumers to show a higher tolerance for them. Some researchers have pointed out that chatbots can mimic human-consumer interactions and create a warm feeling of being understood, valued, and cared for by consumers (Mathies et al., 2016). However, we find no significant mediating effect of perceived warmth. This may be because service failures in the experiment resulted in high levels of negative emotions; further, the eventual outcome was not helpful to consumers in solving a substantial problem. Consequently, consumers did not perceive warmth from either type of service agent.

5.2. Implications

Given the limited research on service failure chatbots, our study is particularly important as it provides insights into how service agent types can help us in understanding consumer attitudes of tolerance after service failure through the perceptual dimensions.

This study makes substantial theoretical contributions. First, it explores the variation in consumers' responses to service failure by investigating their reactions to the negative outcomes of human agents versus Chatbots. While most studies simply compared consumer responses to error and error-free bots (i.e., bot versus bot comparisons), we extend the literature by how people respond to negative outcomes when dealing with humans versus bots. Second, we reveal a crucial insight: consumers generally perceive human agents as more competent and warmer than chatbots when confronted with similar service outcomes. This enriches the human-robot interaction literature, as we show whether people are likely to engage with bots the same way they do with humans (Ho et al., 2018; Mou et al., 2019). Finally, we show that perceived competence, rather than perceived warmth, plays a mediating role in the relationship between service agent type and tolerance in the case of service failure. This may be because service competence is frequently seen as a central factor in determining service level (Boshoff & Allen, 2000). Consumers are more inclined to pursue goal-oriented services, and therefore, are more concerned with the service agents' ability to solve problems rather than the amount of emotional support they provide to customers. This furthers our knowledge on perceived competence as part of the evaluation of service agents, and in gaining a more in-depth understanding of downstream consumer responses and behavioral patterns (Nguyen, 2016; Singh & Kaur, 2011; Wu et al., 2015).

This study also has profound practical value for future chatbots applications. First, considering the current level of AI technology development, the failure of chatbots services is almost unavoidable and requires companies to thoroughly understand how consumers interact with service agents. Recent research shows that the position of human employees in the minds of consumers during service delivery interactions cannot be ignored. While the trend of automating online customer service using chatbots is gaining momentum, the empathy and abilities of human agents remain superior to those of artificial chatbots. Clearly, the human-robot cooperation model needs to be strengthened to improve the quality and efficiency of services.

Second, to further enhance customer satisfaction, service providers should pay attention to the ability of service agents, especially chatbots, to respond to and resolve the situation when confronted with service failures. In particular, the recognition of consumers' emotions should be enhanced and chatbots are able to provide feedback independently and quickly. This can increase consumers' comfort and perception of competence, and reduce the negative effects from AI service failures.

Finally, companies providing AI-based services should not only focus on remedial measures after failure but also on preventive recovery efforts. To reduce the potential negative consequences of AI service failures, companies should disseminate information on preventive measures. Relatively less dangerous information should be disclosed before a potential service failure occurs so that consumers can build their mental protection earlier when confronted with such risks in the future.

5.3. Limitations and Future Research

First, the generalizability of our findings is limited, as a relatively small sample of 119 university students' perceptions of and opinions of service chatbots in sales-oriented consumer service settings. Further studies require a more diverse and broader sample, along with other consumer service settings and kinds of AI. Second, this study only investigated the impact of two service agent types on customer tolerance under service failure scenarios. Scholars can also examine consumers' willingness to use the service under other conditions, especially the effect of a decline in trust. Finally, the chatbots in this research were virtual agents operated with the help of a platform and were not real-time intelligent customer service agents. Therefore, the chatbots in this experiment would have looked more human-like. Further research is needed to test more real-life intelligent assistants. As an increasing number of artificial chatbot applications will be used in various service industries in the immediate future, field experiments are recommended to gather field experience data and provide more comprehensive results on the use of AI technology.

References

Aaker, J., Vohs, K. D., & Mogilner, C. (2010). Nonprofits are seen as warm and for-profits as competent: Firm stereotypes matter. *Journal of Consumer Research, 37*(2), 224-237.

Augusto de Matos, C., Vargas Rossi, C. A., Teixeira Veiga, R., & Afonso Vieira, V. (2009). Consumer reaction to service failure and recovery: the moderating role of attitude toward complaining. *Journal of Services Marketing, 23*(7), 462-475.

Babel, F., Kraus, J., Miller, L., Kraus, M., Wagner, N., Minker, W., & Baumann, M. (2021). Small talk with a robot? The impact of dialog content, talk initiative, and gaze behavior of a social robot on trust, acceptance, and proximity. *International Journal of Social Robotics*, *13*, 1485-1498.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.

Batra, D., & Antony, S. R. (2001). Consulting support during conceptual database design in the presence of redundancy in requirements specifications: an empirical study. *International Journal of Human-Computer Studies, 54*(1), 25-51.

Belanche, D., Casaló, L. V., & Flavián, C. (2020). Customer's acceptance of humanoid robots in services: the moderating role of risk aversion. *Marketing and Smart Technologies: Proceedings of ICMarkTech 2019* (pp.449-458). November 29, Springer, Singapore.

Belanche, D., Casaló, L. V., Schepers, J., & Flavián, C. (2021). Examining the effects of robots' physical appearance, warmth, and competence in frontline services: The Humanness-Value-Loyalty model. *Psychology & Marketing, 38*(12), 2357-2376.

Bergmann, K., Eyssel, F., & Kopp, S. (2012). A second chance to make a first impression? How appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time. *Intelligent Virtual Agents: 12th International Conference, IVA 2012* (pp.126-128). September 12-14, Santa Cruz, CA, USA.

Boshoff, C., & Allen, J. (2000). The influence of selected antecedents on frontline staff's perceptions of service recovery performance. *International Journal of Service Industry Management, 11*(1), 63-90.

Chi, O. H., Denton, G., & Gursoy, D. (2020). Artificially intelligent device use in service delivery: A systematic review, synthesis, and research agenda. *Journal of Hospitality Marketing & Management, 29*(7), 757-786.

Chi, O. H., Gursoy, D., & Chi, C. G. (2022). Tourists' attitudes toward the use of artificially intelligent (AI) devices in tourism service delivery: moderating role of service value seeking. *Journal of Travel Research, 61*(1), 170-185.

Choi, S., & Mattila, A. S. (2008). Perceived controllability and service expectations: Influences on customer reactions following service failure. *Journal of Business Research, 61*(1), 24-30.

Choi, S., Mattila, A. S., & Bolton, L. E. (2021). To err is human (-oid): how do consumers react to robot service failure and recovery? *Journal of Service Research, 24*(3), 354-371.

Crisafulli, B., & Singh, J. (2017). Service failures in e-retailing: Examining the effects of response time, compensation, and service criticality. *Computers in Human Behavior, 77*, 413-424.

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.

Cuddy, A. J., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology, 40*, 61-149.

Fisher, R. A. (1919). XV.—The correlation between relatives on the supposition of Mendelian inheritance. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh, 52*(2), 399-433.

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77-83.

Følstad, A., & Brandtzæg, P. B. (2017). Chatbots and the new world of HCI. *Interactions, 24*(4), 38-42.

Følstad, A., Nordheim, C. B., & Bjørkli, C. A. (2018). What makes users trust a chatbot for customer service? An exploratory interview study. *Internet Science: 5th International Conference, INSCI 2018* (pp. 194-208). October 24-26, St. Petersburg, Russia.

Gailey, J. A. (2013). Attribution of responsibility for organizational wrongdoing: A partial test of an integrated model. *Journal of Criminology, 2013*, 920484.

Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition, 125*(1), 125-130.

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry, 23*(2), 101-124.

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling. University of Kansas, KS.

Heider, F. (2013). *The psychology of interpersonal relations.* Psychology Press.

Ho, A., Hancock, J., & Miner, A. S. (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication, 68*(4), 712-733.

Hong, J.-W., & Williams, D. (2019). Racism, responsibility and autonomy in HCI: Testing perceptions of an AI agent. *Computers in Human Behavior, 100,* 79-84.

Kaczorowska-Spychalska, D. (2019). How chatbots influence marketing. *Management, 23*(1), *251-270.*

Keaveney, S. M. (1995). Customer switching behavior in service industries: An exploratory study. *Journal of Marketing, 59*(2), 71-82.

Kim, S. Y., Schmitt, B. H., & Thalmann, N. M. (2019). Eliza in the uncanny valley: Anthropomorphizing consumer robots increases their perceived warmth but decreases liking. *Marketing Letters, 30,* 1-12.

Kirmani, A., Hamilton, R. W., Thompson, D. V., & Lantzy, S. (2017). Doing well versus doing good: The differential effect of underdog positioning on moral and competent service providers. *Journal of Marketing, 81*(1), 103-117.

Lee, K. M., Jung, Y., Kim, J., & Kim, S. R. (2006). Are physically embodied social agents better than disembodied social agents? The effects of physical embodiment, tactile interaction, and people's loneliness in human–robot interaction. *International Journal of Human-Computer Studies, 64*(10), 962-973.

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.

Liu, B., & Li, Y. (2022). Teddy-bear effect in service recovery. *Annals of Tourism Research, 94,* 103400.

Liu, P., Glas, D. F., Kanda, T., & Ishiguro, H. (2018). Learning proactive behavior for interactive social robots. *Autonomous Robots, 42,* 1067-1085.

Lou, C., Kang, H., & Tse, C. H. (2022). Bots vs. humans: how schema congruity, contingency-based interactivity, and sympathy influence consumer perceptions and patronage intentions. *International Journal of Advertising, 41*(4), 655-684.

Lu, L., Zhang, P., & Zhang, T. C. (2021). Leveraging "human-likeness" of robotic service at restaurants. *International Journal of Hospitality Management, 94,* 102823.

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, 38*(6), 937-947.

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.

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.

Mathies, C., Chiew, T. M., & Kleinaltenkamp, M. (2016). The antecedents and consequences of humour for service: A review and directions for research. *Journal of Service Theory and Practice, 26*(2), 137-162.

Mieczkowski, H., Liu, S. X., Hancock, J., & Reeves, B. (2019). Helping not hurting: Applying the stereotype content model and bias map to social robotics. *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 222-229). March 11-14, Daegu, Korea (South).

Mou, Y., Xu, K., & Xia, K. (2019). Unpacking the black box: Examining the (de) Gender categorization effect in human-machine communication. *Computers in Human Behavior, 90,* 380-387.

Newton, J. D., Wong, J., & Casidy, R. (2018). Deck the halls with boughs of holly to soften evaluations of service failure. *Journal of Service Research, 21*(4), 389-404.

Nguyen, N. (2016). Reinforcing customer loyalty through service employees' competence and benevolence. *The Service Industries Journal, 36*(13-14), 721-738.

Oliveira, R., Arriaga, P., Correia, F., & Paiva, A. (2019, March). The stereotype content model applied to human-robot interactions in groups. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 123-132). March 11-14, Daegu, Korea (South).

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.

Pitardi, V., & Marriott, H. R. (2021). Alexa, she's not human but… Unveiling the drivers of consumers' trust in voice-based artificial intelligence. *Psychology & Marketing, 38*(4), 626-642.

Rajaobelina, L., Prom Tep, S., Arcand, M., & Ricard, L. (2021). Creepiness: Its antecedents and impact on loyalty when interacting with a chatbot. *Psychology & Marketing, 38*(12), 2339-2356.

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places, first ed,* Cambridge University Press.

Sheehan, B., Jin, H. S., & Gottlieb, U. (2020). Customer service chatbots: Anthropomorphism and adoption. *Journal of Business Research, 115,* 14-24.

Singh, J., & Kaur, G. (2011). Customer satisfaction and universal banks: an empirical study. *International Journal of Commerce and Management, 21*(4), 327-348.

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*(3), 356-372.

Sparks, B. A., & Browning, V. (2010). Complaining in cyberspace: The motives and forms of hotel guests' complaints online. *Journal of Hospitality Marketing & Management, 19*(7), 797-818.

Sun, Y. (2021). Case based models of the relationship between consumer resistance to innovation and customer churn. *Journal of Retailing and Consumer Services, 61,* 102530.

Townsend, C., & Sood, S. (2012). Self-affirmation through the choice of highly aesthetic products. *Journal of Consumer Research, 39*(2), 415-428.

Tsarenko, Y., Strizhakova, Y., & Otnes, C. C. (2019). Reclaiming the future: Understanding customer forgiveness of service transgressions. *Journal of Service Research, 22*(2), 139-155.

Tussyadiah, I. (2020). A review of research into automation in tourism: Launching the Annals of Tourism Research Curated Collection on Artificial Intelligence and Robotics in Tourism. *Annals of Tourism Research, 81,* 102883.

Wang, X., & Zhang, Q. (2018). Does online service failure matter to offline customer loyalty in the integrated multi-channel context? The moderating effect of brand strength. *Journal of Service Theory and Practice, 28*(6), 774-806.

Wang, Z., Mao, H., Li, Y. J., & Liu, F. (2017). Smile big or not? Effects of smile intensity on perceptions of warmth and competence. *Journal of Consumer Research, 43*(5), 787-805.

Weiner, B. (2000). Attributional thoughts about consumer behavior. *Journal of Consumer Research, 27*(3), 382-387.

Wilson, H. J., & Daugherty, P. R. (2018). Collaborative intelligence: Humans and AI are joining forces. *Harvard Business Review, 96*(4), 114-123.

Wojciszke, B., Abele, A. E., & Baryla, W. (2009). Two dimensions of interpersonal attitudes: Liking depends on communion, respect depends on agency. *European Journal of Social Psychology, 39*(6), 973-990.

Wortman, J., & Wood, D. (2011). The personality traits of liked people. Journal of Research in Personality, 45(6), 519-528.

Wu, Y.-C., Tsai, C.-S., Hsiung, H.-W., & Chen, K.-Y. (2015). Linkage between frontline employee service competence scale and customer perceptions of service quality. *Journal of Services Marketing, 29*(3), 224-234.

Xu, Y., Niu, N., & Zhao, Z. (2023). Dissecting the mixed effects of human-customer service chatbot interaction on customer satisfaction: An explanation from temporal and conversational cues. *Journal of Retailing and Consumer Services, 74,* 103417.