INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES

Vol 15, Issue 11, (2025) E-ISSN: 2222-6990

The Impact of AI Customer Service Incivility on Consumer Service Experience in Retail E-Commerce Platforms: An E-Politeness Perspective

I-Ching Chen*, Li-Qian Quan

School of Economics and Management, Zhaoqing University, 526060, Zhaoqing, China

*Corresponding Author Email: jineandya@gmail.com

DOI Link: http://dx.doi.org/10.6007/IJARBSS/v15-i11/26944

Published Date: 28 November 2025

Abstract

While AI customer service has effectively contributed to cost reduction and operational efficiency on retail e-commerce platforms, its interactional shortcomings—particularly in the form of "incivility"—have significantly undermined consumer service experiences. Adopting a service politeness perspective, this study systematically investigates how uncivil behaviors exhibited by AI agents trigger negative consumer responses. Using the Critical Incident Technique, 325 valid negative incidents were collected from major Chinese social media platforms and subjected to qualitative analysis. Findings reveal six core dimensions of uncivil AI behavior: problem-solving efficacy, response accuracy, human handover, empathetic engagement, system stability, and reliability of commitment. Among these, issues related to human handover (29.23%) and response accuracy (26.15%) were the most prevalent. This study not only fills a theoretical gap in the politeness dimension of AI service experience but also provides actionable insights and empirical support for optimizing human–machine interaction design in digital commerce. Based on the findings, targeted recommendations are proposed for four key stakeholders—platform governance, technical functionality, merchants, and consumers—to collectively enhance the quality and civility of AI-driven customer service.

Keywords: AI Customer Service, Retail E-Commerce Platforms, E-Politeness, Service Experience, Critical Incident Technique

Introduction

As the scale of e-commerce continues to expand and market competition intensifies, cost reduction and service experience enhancement have become core imperatives for platform operators. Against this backdrop, AI-powered customer service has been widely adopted in retail e-commerce due to its notable advantages in reducing labor costs, enabling 24/7 instant responsiveness, and standardizing the handling of large volumes of frequent inquiries. However, technological empowerment also brings forth new challenges. Current AI customer service systems predominantly rely on pre-defined corpora and pattern-matching algorithms, which often struggle to address complex and emotionally nuanced queries. These shortcomings hinder resolution and provoke consumer dissatisfaction. As a real-time communication bridge between firms and consumers, customer service plays a pivotal role in shaping service experience, brand perception, and ultimately purchase decisions. Therefore, the deficiencies observed in AI customer service are not merely technical imperfections but potential threats to customer satisfaction and brand trust. Although both academia and industry have begun to pay attention to the experiential aspects of AI customer service, existing studies predominantly approach the topic from an efficiency-oriented perspective (Pan, 2024; Zhang et al., 2025), with limited exploration into the quality of human–AI interaction from the lens of e-politeness. This theoretical gap constrains platforms' ability to deeply understand and fundamentally improve the service experience embedded in AI-mediated interactions.

To address the aforementioned research gap, this study introduces Whitworth's (2005) theory of computer politeness as the conceptual framework for analyzing politeness in AI-mediated interactions. Employing a qualitative research design grounded in the Critical Incident Technique, the study aims to uncover the underlying mechanisms of politeness deficiency in AI customer service within e-commerce contexts. The findings are intended to offer actionable theoretical insights and strategic recommendations for retail e-commerce platforms seeking to optimize AI-driven service interactions and enhance consumer experience.

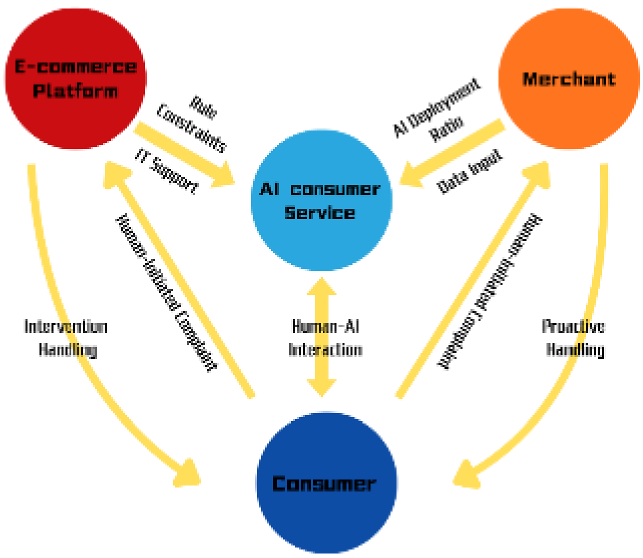

Figure 1. Ecosystem of Stakeholder Interactions in AI-Driven Customer Service on Retail E-Commerce Platforms

Literature Review

Service-Oriented Experience Deficiency and Reconstruction in E-Commerce

Technological advancements—such as AI-driven personalized recommendations, augmented and virtual reality for enhanced product experiences, IoT-enabled inventory optimization, and 5G-powered seamless connectivity—have emerged as key drivers of e-commerce development (Azam & Ansari, 2024). Compared to traditional offline transactions, e-commerce transcends temporal and spatial constraints, allowing consumers to shop anytime and anywhere. This convenience has become a central factor in consumer decision-making (Mustafa et al., 2022). Moreover, e-commerce platforms facilitate rapid price comparisons, enabling consumers to identify high-value products efficiently. The resulting transparency significantly reduces search costs and enhances perceived value in terms of price and information accessibility (Suyanto et al., 2019; Peral Peral et al., 2012), further contributing to the appeal of online shopping. However, despite these efficiency-driven benefits, the e-commerce model exhibits inherent limitations. Unlike offline retail environments, online shopping lacks immediate non-verbal communication, rich multisensory product demonstrations, trust-building and problem-solving—key aspects of the valued 'human touch'. (Grosso & Forza, 2019; Hamacher & Buchkremer, 2022; Komiak & Benbasat, 2004; Ming, 2022; Eusebio et al., 2023). The absence of these experiential elements constitutes a core tension in the development of e-commerce, particularly from the perspective of politeness and interpersonal engagement. To meet consumers' holistic expectations, platforms must actively address these deficiencies. Given that many of the missing experiences stem from the lack of face-to-face interaction, enhancing the quality and experience of online customer service emerges as a critical pathway for resolution. Therefore, enhancing the quality of online customer service becomes a critical pathway to compensate for these missing experiential elements.

From Interpersonal Politeness to Systemic Interaction

For consumers, politeness is not merely a superficial social courtesy but a reflection of deeper psychological needs for respect and fairness—core elements of the psychological contract that underpins commercial exchanges (Goffman, 2017; Maslow, 1943; Rousseau, 1995). In offline retail settings, this contract is conveyed and maintained through various interpersonal cues such as language, facial expressions, and body gestures, which play a vital role in building trust (Karthik & Sreehari, 2025). However, the "faceless" nature of e-commerce interactions fundamentally narrows the scope for expressing politeness, stripping away most social signals inherent in human communication. As a result, traditional politeness behaviors rooted in interpersonal skills must be reconfigured into system-mediated forms embedded within technological design. In response to this shift, Whitworth's (2005) Politeness Theory offers a critical analytical framework, positing that politeness in online systems is manifested through adherence to three fundamental principles: Rights, Goals, and Information. The significance of this framework lies in its ability to transform e-commerce politeness from a vague, subjective emotional experience into a set of clearly definable and measurable system functions. In the absence of non-verbal compensatory cues, the extent to which service providers respect user rights (e.g., privacy and autonomy), support user goals (e.g., seamless checkout processes), and deliver accurate and transparent information (e.g., inventory and logistics updates) directly influences perceived consumer satisfaction (Gelashvili et al., 2024; Ali, 2024; Ortega-Bolaños et al., 2024; Chen & Lu, 2025).In essence, consumer expectations of politeness in retail e-commerce have evolved beyond scripted customer service dialogues. They now demand a systemically designed experience that integrates robust technological and professional support throughout the entire shopping journey—compensating for the emotional engagement that is inherently absent in digital environments.

The Dual Nature of AI Customer Service: Operational Efficiency and Experiential Deficiencies

In recent years, AI customer service has emerged as a pivotal component in the digital transformation of enterprises, evolving from a supplementary tool into an indispensable element of customer service systems (Larivière et al., 2017). This evolution is largely driven by advancements in core technologies such as Natural Language Processing (NLP), Machine Learning (ML), and Automatic Speech Recognition (ASR), which enable automated query handling, sentiment analysis, and personalized responses—substantially enhancing service efficiency and user satisfaction (Pan, 2024). Within the e-commerce sector, AI customer service has become a foundational pillar of the online retail ecosystem, offering instant responsiveness, personalized engagement, and round-the-clock availability (Sahu, 2025; Rasheed et al., 2025). Despite its rapid development, AI customer service continues to face notable challenges. These include the generation of inaccurate or misleading information (commonly referred to as the “hallucination” problem), vulnerabilities in data security and user privacy protection, and a lack of empathy when addressing complex or emotionally charged inquiries (Clelland et al., 2024; Abiagom & Ijomah, 2024; Xie et al., 2025). Such deficiencies may lead to consumer dissatisfaction and perceptions of incivility. Resulting in uncivil service experiences. These negative interactions not only provoke immediate dissatisfaction but may also erode the platform’s service reputation and consumer trust. Consequently, retail e-commerce platforms and merchants must carefully weigh the benefits of AI-driven service innovations against their potential risks, striving to enhance operational efficiency while safeguarding the integrity of the consumer experience (Zhao & Wu, 2025).

Research Methodology

Critical Incident Technique

The Critical Incident Technique (CIT) is a qualitative research method that systematically collects, analyzes, and categorizes observational data on human behavior. It has been widely adopted across disciplines for its ability to capture specific actions within real-life contexts rather than relying on abstract opinions or attitudes (Gremler, 2015). The core of CIT lies in identifying and examining events that exert significant influence within a particular setting (Jombe & Pretorius, 2025). By focusing on the incident itself, researchers can gain deep insights into its causes, processes, and outcomes, thereby uncovering underlying mechanisms and contributing factors (Jombe & Pretorius, 2025). Originally introduced by Flanagan (1954) in the social sciences, CIT has evolved over more than six decades and found applications in various fields. In education, Rotem et al. (2024) employed CIT to analyze critical moments in teaching practice, facilitating reflective discussions among pre-service teachers during field-based university courses and enhancing their attentional awareness. In hospitality management, Baker and Kim (2019) used CIT to explore the motivations and perceptions of exaggerated reviews, revealing both micro-level generation mechanisms and macro-level value-destruction effects. In psychology, Durosini et al. (2021) applied CIT to investigate individuals’ experiences during COVID-19 lockdowns, identifying a unique sense of urgency in emotional responses to isolation-related stimuli. In management studies, Madsen and Ulhøi (2021) utilized CIT to identify pivotal transformation points in corporate change processes, demonstrating that a sustainable vision can provide both urgency and strategic direction for innovation and long-term development.

In summary, CIT has matured into a robust methodology across multiple disciplines. Given this study’s aim to explore uncivil incidents triggered by AI customer service on retail e-commerce platforms—focusing on context-specific behaviors and reactions, and uncovering their underlying causes and impact mechanisms—CIT is adopted as the primary research method.

Data Collection

Given the nature of this study, negative service experiences tend to leave consumers with deeper and more vivid memories compared to neutral or positive ones, often prompting detailed narratives and public sharing on social platforms (Bitner et al., 1990). These negative incidents serve as critical diagnostic tools, revealing specific breakdowns and design flaws in AI customer service interactions, as well as their emotional and behavioral consequences. Such incidents offer rich and direct evidence for investigating the causes and outcomes of uncivil service experiences. The data for this study were sourced from publicly available consumer posts on three major Chinese social media platforms: Sina Weibo, Douyin, and Xiaohongshu. These platforms are widely used by consumers to vent frustrations, seek empathy, and lodge public complaints following service failures. Their open content naturally forms a data pool centered on negative experiences, effectively aligning with the study's focus on incivility in AI-mediated service encounters. Data collection was conducted between September 2 and 7, 2025. Using keyword searches such as "AI customer service", "Intelligent customer service", in combination with terms reflecting negative sentiment like "rant", "angry", and "impolite", a large volume of relevant posts was initially retrieved. After a preliminary review, posts and comments that clearly described negative experiences during interactions with AI customer service on retail e-commerce platforms were selected, resulting in an initial dataset of 847 valid incident descriptions. This dataset laid the foundation for subsequent rigorous screening and qualitative analysis.

Data Analysis

Descriptive Overview

Through manual data collection conducted by the research team on three major Chinese social media platforms—Sina Weibo, Douyin, and Xiaohongshu—a total of 847 incident reports were initially gathered. To ensure user privacy, all cases were anonymized prior to further screening. A two-stage filtering process was employed. In the first stage, irrelevant or noisy data were quickly excluded based on thematic relevance, such as non-event narratives, incorrect subjects, unrelated platforms, and duplicate content. The remaining cases proceeded to a second stage of in-depth screening. At this stage, two researchers independently reviewed the cases using four inclusion criteria: (1) Completeness—the incident must contain a full interaction context, process, and outcome; (2) Relevance—the event must involve uncivil behavior by AI customer service on retail e-commerce platforms; (3) Specificity—the description must include concrete details or dialogue rather than vague generalizations; and (4) Representativeness—the incident should typify a pattern of incivility or have a significant impact on the consumer. Discrepancies in judgment were resolved through discussion until consensus was reached. Following this rigorous screening, a final set of 325 valid critical incidents involving uncivil AI customer service interactions on retail e-commerce platforms was retained. These cases form the analytical foundation for subsequent categorization and interpretation. Detailed information on the final dataset is presented in Table 1.

Table 1
Summary of Critical Incident Data Collection

Platform	Time Frame	Initial Cases Collected	Valid Incidents After Screening	Proportion
Sina Weibo	2024.08-2025.08	211	67	20.62%
Douyin		299	112	34.46%
Xiaohongshu		339	146	44.92%
Total	-	847	325	100%

Classification Principles

To ensure the reliability of research findings, this study employed a systematic procedure for categorizing critical incidents. First, two researchers independently conducted close readings and content analyses of the screened incidents. Through iterative discussions, they aimed to derive mutually exclusive and collectively exhaustive categories that could encompass all observed cases. Ultimately, six categories were identified: problem-solving efficacy, response accuracy, human handover, empathetic engagement, system stability, and reliability of commitment (see Table 2 for category definitions). Once the category framework was finalized, three experienced coders—who had not participated in the initial framework development—were invited to perform two rounds of independent classification. Prior to coding, the researchers provided a detailed definition of "incivility" to guide coders in identifying and assigning incidents appropriately. In the first round, the anonymized and randomized incident dataset was distributed to the coders, who classified each case based on the definitions in Table 2. One month later, the coders repeated the classification process with the same dataset to assess consistency. Background information on the three coders involved in this task is presented in Table 3.

Table 2
Category Labels and Definitions of Critical Incidents

Category Label	Definition
Problem-Solving Efficacy	The efficiency and effectiveness in receiving, processing, and resolving consumer's specific requests.
Response Accuracy	The core capability to generate accurate replies after identifying consumer intent and understanding natural language queries.
Human Handover	The entire experience from initiating a request to transfer to a human agent to either successful connection or final failure.
Empathetic Engagement	The consumer's subjective emotional experience during interaction with the customer service system.
System Stability	The reliability and robustness of the backend systems and technical infrastructure supporting consumer service interactions.
Reliability of Commitment	The accuracy of promises made by the AI customer service and the extent to which they are fulfilled.

Table 3
Background Information of Coders

Coder	Occupation	Background Summary
Coder A	Associate Professor in E-Commerce	Over five years of teaching experience in e-commerce; strong theoretical foundation and academic expertise.
Coder B	Human Customer Service Agent	Seven years of experience in e-commerce customer service; well-versed in platform operations and user needs.
Coder C	AI Customer Service Developer	Three years of programming experience; involved in multiple AI customer service system development projects with deep understanding of backend logic.

Reliability and Validity Analysis

Reliability Assessment

Reliability refers to the consistency and stability of a measurement instrument when applied repeatedly under similar conditions (Artino et al., 2014). In this study, reliability analysis was conducted by comparing the results of two rounds of incident classification, with the aim of ensuring high inter-coder agreement in identifying, categorizing, and evaluating critical incidents. It also serves to confirm that the incident descriptions themselves reliably and accurately reflect real-world experiences. Prior research suggests that a reliability coefficient of 0.80 or above indicates strong consistency and provides a robust foundation for subsequent analysis (Flanagan, 1954). In this study, the classification results from three coders across two rounds were analyzed to determine the number of consistently coded incidents. The results are summarized in Table 4.

Table 4
Inter-Coder Consistency—Number of Consistently Classified Uncivil AI Customer Service Incidents

	Coder A	Coder B	Coder C
Coder A	296	-	-
Coder B	192	307	-
Coder C	223	216	284

Based on the data presented in Table 4, this study conducted a reliability test on the classification results of the three coders. The objective was to evaluate the degree of consistency across coders in categorizing uncivil AI customer service incidents. The reliability coefficient was calculated using the following formula:

$$A=\frac{\frac{2M_{12}}{n_1+n_2}+\frac{2M_{23}}{n_2+n_3}+\frac{2M_{13}}{n_1+n_3}}{N} \quad (1)$$

$$R=\frac{(N\times A)}{1+[(N-1)\times A]} \quad (2)$$

Where: R= Reliability

N= Number of coders

A= Average interjudge agreement

M= Number of consistently classified incidents between coders (e.g., M_{12} refers to the number of incidents classified identically by Coder 1 and Coder 2)

n= Number of classification judgments made by each coder (e.g.,n_1 refers to the total number of incidents classified by Coder 1)

Based on the above formula and subsequent calculations, the final reliability coefficient was determined to be 0.881. This result demonstrates strong inter-coder reliability in the classification process, confirming that the dataset used in this study is methodologically sound and capable of supporting further analysis with confidence.

Validity Assessment

Validity refers to the extent to which a research instrument, measurement method, or study design accurately captures the intended construct, ensuring the authenticity, precision, and applicability of the findings (Ansari & Khan, 2023; Hirano & Roberson, 2010). Validity is commonly categorized into content validity, face validity, and construct validity (Lim, 2024). This study primarily evaluates the classification system of critical incidents through content and construct validity. Content validity assesses whether the measurement tool adequately and accurately covers all key dimensions of the concept or behavior it intends to measure. To ensure content validity, this study followed a structured approach: first, the operational definition of “uncivil experience” was established based on Whitworth’s (2005) theory of computer politeness. Then, two researchers independently reviewed and categorized all critical incidents according to this definition. Full consensus was reached on the final six category labels, confirming that the classification system comprehensively captured the essential dimensions of uncivil AI customer service experiences (Ansari & Khan, 2023). Construct validity refers to the degree to which a measurement tool accurately reflects the theoretical construct it claims to assess. In this study, construct validity was ensured through a systematic classification procedure in which all incidents were independently coded by three trained coders. This process minimized subjective bias and ensured that the final category structure emerged directly from the data, authentically representing the underlying dimensions of the “uncivil experience” construct (Olanipekun et al., 2022).

Classification Results

Following the completion of incident categorization and reliability testing, this study proceeded to a deeper analysis of how uncivil AI customer service interactions affect consumer experiences on retail e-commerce platforms. To illustrate the nature and impact of these interactions, two representative incidents were selected from each of the six identified categories. These examples serve to highlight typical patterns of uncivil behavior and their consequences. Detailed descriptions of the selected critical incidents are presented in Table 5.

Table 5
Representative Critical Incident Examples by Category

Category Label	Representative Critical Incidents
Problem-Solving Efficacy	I asked it to issue an invoice, but it kept going in circles. This solution didn't work, that one didn't either—nothing was resolved.
	I bought a tire and later found the price had dropped. I requested a price guarantee, and the platform said a specialist would contact me. When I followed up, the AI said it had urged them—but after 7–8 days, no one reached out.
Response Accuracy	I asked whether return shipping was covered, but it kept replying with scripted phrases like 'We understand your urgency'.
	I asked how to install the product, but it kept giving irrelevant answers, as if it didn't see my question.
Human Handover	I tried to reach a human agent but couldn't. Then I cursed 'artificial idiot,' and suddenly the human service window popped up.
	I've been trying for two days to get a human agent. All I get are replies from the Taoxiaomi bot, repeating the same lines.
Empathetic Engagement	It just kept saying 'We understand your feelings, we're sorry.' Honestly, I felt completely misunderstood.
	So frustrating—felt like talking to thin air. Just repeated apologies, no solutions, no emotional understanding.
System Stability	The handover button appeared, but after ten minutes of queuing, the time didn't decrease. When I came back, the session had ended automatically.
	During the Double 11 sale, I asked about logistics, but the AI seemed to crash. It couldn't recognize my question and just recommended products I might like.
Reliability of Commitment	I had a broken item and contacted customer service. The AI said it was recorded and someone would follow up—but no one ever did.
	I bought a thermostat on Taobao. The seller didn't ship on time, and the AI promised 30% compensation. But later, it turned into a 30-yuan coupon with no explanation.

The representative critical incidents presented above reveal a consistent pattern: consumers enter interactions with clear and specific requests, yet AI customer service systems repeatedly fail to address these core needs. This failure not only deteriorates the service experience but also triggers perceptions of incivility. Viewed across the entire interaction process, AI customer service exhibits a fundamental deficiency in problem-solving capabilities. Consumers often find themselves trapped in an "AI maze," unable to reach human agents, and simultaneously deprived of empathetic understanding. This leads to a dual deficiency—neither functional resolution nor emotional support is provided. Over time, such recurring deficiencies may lead to a systemic erosion of consumer trust. When platforms fail to deliver both instrumental and relational value, the perceived reliability of the service infrastructure is undermined, posing significant risks to brand credibility and long-term customer loyalty.

Subsequently, this study conducted statistics and analysis on the key events after classification. The specific data are shown in Table 6 below.

Table 6
Statistics of Event Category Data

Category Label	Number of Incidents	Proportion
Problem-Solving Efficacy	70	21.45%
Response Accuracy	85	26.15%
Human Handover	95	29.23%
Empathetic Engagement	40	12.31%
System Stability	20	6.15%
Reliability of Commitment	15	4.62%

As shown in Table 6, issues related to human handover account for the largest proportion of critical incidents, nearly 30%. This is followed by response accuracy, which constitutes 26.15%. Together, these two categories represent more than half of all uncivil AI customer service experiences reported by consumers. This finding highlights a strategic priority for retail e-commerce platforms: addressing deficiencies in AI-driven handover mechanisms and improving the accuracy of automated responses can substantially reduce the occurrence of uncivil service experiences. By resolving these two core issues, platforms can significantly enhance the overall quality of customer service and foster more positive consumer interactions.

Conclusion and Recommendations

Conclusion

Drawing on service politeness theory, this study analyzes 325 negative AI service experiences from Weibo, Xiaohongshu, and Douyin.. Using the Critical Incident Technique, 325 valid incidents were qualitatively examined. The findings reveal six core dimensions of uncivil AI behavior: problem-solving efficacy, response accuracy, human handover, empathetic engagement, system stability, and reliability of commitment. Among these, human handover (29.23%) and response accuracy (26.15%) emerged as the most prevalent sources of dissatisfaction. These results suggest that consumers frequently encounter AI systems that fail to accurately understand or resolve their requests, while also facing barriers when attempting to escalate issues to human agents. In-depth analysis of representative incidents further reveals that uncivil AI service experiences often stem from a dual failure: the inability to solve problems and the absence of emotional support. These dual failures leave consumers feeling misunderstood and unsupported. These issues stem more from governance ambiguity and system flaws than from AI itself. To fully realize the cost-efficiency potential of AI customer service, retail e-commerce platforms must move beyond a narrow focus on operational efficiency and adopt a holistic politeness-oriented approach to system design. Accordingly, the study proposes targeted recommendations for platforms, systems, merchants, and consumers to improve AI service civility.

This study has two main limitations. First, all data were sourced from Chinese social media, which enables in-depth analysis but limits generalizability due to China's unique digital ecosystem and consumer culture. Second, the focus on retail e-commerce may restrict applicability to other service sectors with different processes and pain points. Future research should pursue two directions: cross-cultural comparisons to test the framework's relevance in varied markets, and expansion into other service industries to identify both shared and sector-specific patterns. These efforts will deepen understanding of AI–consumer interactions and inform more adaptive service strategies.

Theoretical and Contextual Contributions

This study provides significant theoretical and contextual contributions. Theoretically, it extends Whitworth's (2005) politeness theory into the novel domain of AI-mediated service interactions, moving beyond the traditional human-to-human paradigm. By identifying and validating six core dimensions of AI incivility, it offers a structured conceptual framework that enriches the service experience literature, which has historically been dominated by an efficiency-oriented perspective. This bridges a critical theoretical gap and provides a measurable foundation for future research on AI-consumer rapport.

Contextually, the findings offer critical insights for the rapidly evolving Chinese retail e-commerce landscape. They reveal how systemic failures in AI design—particularly in human handover and response accuracy—are perceived as uncivil behaviors, undermining consumer trust in a high-stakes, high-volume market. This contextual understanding equips platform operators, merchants, and developers with empirically grounded guidance to optimize AI service design, not only to mitigate negative experiences but also to foster a more polite, resilient, and trustworthy human-AI service ecosystem. The study thus serves as a strategic tool for enhancing service quality in practice while advancing academic knowledge in the field of digital service interactions.

Suggestion

Rule-Based Recommendations for Platform Governance

Platform rules form the top-level framework for regulating AI service and safeguarding consumer experience. Based on the study's findings, three recommendations are proposed to enhance the perceived politeness of platform governance: First, clearly define the service boundaries between AI and human customer service. In response to issues related to "human handover" and "problem-solving efficacy," Platforms should move beyond treating AI as a gatekeeping tool and clearly delineate service boundaries. If an issue remains unresolved, the system should promptly transfer the case to a human agent for resolution. Second, establish a commitment tracking protocol for AI customer service. To address deficiencies in "reliability of commitment," platforms should implement a rule that any promise made by the AI must be fulfilled. All commitments involving follow-up actions (e.g., "You will receive a reply within 24 hours," "The coupon will be issued") must be systematically recorded and tracked to ensure execution. This mechanism helps prevent trust erosion caused by unfulfilled promises and enhances the platform's credibility. Third, implement a routine inspection mechanism. Platforms should regularly analyze high-frequency complaint cases to identify rule gaps and service blind spots. These insights should inform rule refinement and technical upgrades. A data-driven feedback loop enables platforms to refine rules and improve AI service civility and quality.

Function-Oriented Recommendations for Technical Enhancement

Advanced functional technologies are essential for improving the quality of AI–consumer interactions and upholding service politeness in e-commerce. Based on the study's findings, four recommendations are proposed: First, enable seamless context-preserving handovers. Information discontinuity between AI and human agents undermines problem-solving.

Systems should retain conversation history to avoid repetition, ensuring continuity and professionalism. Second, Develop a data-driven corpus and iterative functionality. To improve response accuracy, platforms should use high-frequency complaint data via routine audits to update the AI's language base and refine functions. This improves real-world problem-solving and reduces ineffective interactions. Third, implement emotion recognition and feedback mechanisms. Real-time sentiment monitoring can identify dissatisfaction triggers. These insights should inform continuous refinement of dialogue and system responses. Fourth, reinforce system architecture for stability and resilience. During peak periods such as major promotional events (e.g., Double 11), platforms must ensure the stable operation of AI customer service systems. Robust infrastructure is critical to maintaining smooth performance and timely responses under high concurrency conditions.

Merchant-Oriented Recommendations for Service Optimization

In the practical operation of AI customer service systems, merchants serve as the direct providers of service. Based on the research findings, three recommendations are proposed to enhance the overall service experience: First, use AI customer service judiciously. Merchants should have a clear understanding of the functional boundaries of AI systems and allocate AI and human agents appropriately. For high-value products, those with high complaint risks, or those involving complex after-sales processes, human customer service should be prioritized to prevent dissatisfaction caused by the AI's limited problem-solving capabilities. Second, maintain a dynamic knowledge base and update mechanism. The "response accuracy" of AI customer service is highly dependent on the quality and timeliness of its backend knowledge base. During key operational moments—such as new product launches, promotional campaigns, or policy changes—merchants must ensure that relevant information is promptly updated to maintain service relevance and accuracy. Third, conduct routine audits of AI service data. Merchants should not allow AI systems to operate unattended after deployment. Instead, they should regularly review conversation logs, with particular attention to sessions with high handover rates, abrupt terminations, or evident consumer dissatisfaction. These records should be systematically analyzed to identify weaknesses and guide targeted improvements.

Consumer-Oriented Recommendations for Co-Creating Service Quality

Within the AI customer service ecosystem, consumers are not only recipients of service but also active participants in shaping service quality. Based on the research findings, two recommendations are proposed: First, strengthen awareness of service feedback. When encountering uncivil AI service experiences, consumers should retain relevant interaction records and report them through official channels provided by the platform or merchant. Timely and constructive feedback is essential for driving system improvements and holding service providers accountable. Second, develop effective human–machine interaction skills. As AI customer service becomes increasingly prevalent, consumers are encouraged to proactively learn and apply key communication strategies. For example, using concise, clear keywords and declarative sentences can improve AI comprehension. If the AI enters a repetitive response loop, consumers should promptly switch service paths or request human assistance. By enhancing their own interaction capabilities, consumers can improve problem-solving efficiency and achieve smoother service experiences—even within the constraints of current technology.

Acknowledgments

This research work was funded by the grant from the Guangdong Science and Technology Program (China) under Grant No. 2024A0505050036, and the grants from the Department of Education of Guangdong Province under Grant Nos: 2021WTSCX093 and 2020GXJK168. We deeply appreciate their financial support and encouragement.

References

Pan, Y. (2024). Research on the Current Status and Development Trends of AI in Customer Service Systems. *Academic Journal of Science and Technology, 13*(3), 74-77. https://doi.org/10.54097/edyb0f32

Zhang, Y., Ding, Z., Sun, J., Zhao, X., Hu, X., & Yang, Z. (2025). Optimizing AI strategies in e-commerce customer service: An agent-based simulation. *Electronic Markets, 35*(1), 1-22. https://doi.org/10.1007/s12525-025-00821-8

Whitworth, B. (2005). Polite computing. *Behaviour & Information Technology, 24*(5), 353-363. https://doi.org/10.1080/01449290512331333700

Azam, A., & Ansari, A. M. (2024). The emerging role of e-commerce in today's business: A conceptual study. *Asian Journal of Management and Commerce, 5*(1), 428-439. https://doi.org/10.22271/27084515.2024.v5.i1f.289

Mustafa, S., Hao, T., Qiao, Y., Shah, K., & Sun, R. (2022). How a successful implementation and sustainable growth of e-commerce can be achieved in developing countries; a pathway towards green economy. *Frontiers in Environmental Science*, *10*, 1-17. https://doi.org/10.3389/fenvs.2022.940659

Suyanto, B., Subiakto, H., & Srimulyo, K. (2019). Data of the patterns of youth local brand product consumption through online shopping. *Data in brief*, *23*, 103723. https://doi.org/10.1016/j.dib.2019.103723

Peral Peral, B., Rey, J. R. B., & Villarejo Ramos, Á. F. (2012). A study of consumer preferences for e-retailer'attributes: An application of conjoint analysis. *The International Journal of Management Science and Information Technology*, *3*(Jan-Mar), 38-62. https://hdl.handle.net/10419/97875

Grosso, C., & Forza, C. (2019). Users' social-interaction needs while shopping via online sales configurators. *International Journal of Industrial Engineering and Management*, *10*(2), 139-154. https://doi.org/10.24867/ijiem-2019-2-235

Hamacher, K., & Buchkremer, R. (2022). Measuring Online Sensory Consumer Experience: Introducing the Online Sensory Marketing Index (OSMI) as a Structural Modeling Approach. *Journal of Theoretical and Applied Electronic Commerce Research, 17*(2), 751-772. https://doi.org/10.3390/jtaer17020039

Komiak, S. X., & Benbasat, I. (2004). Understanding Customer Trust in Agent-Mediated Electronic Commerce, Web-Mediated Electronic Commerce, and Traditional Commerce. *Information Technology and Management, 5*(1-2), 181-207. https://doi.org/10.1023/b:item.0000008081.55563.d4

Ming, G. (2022). Online retail service quality, Who satisfy? - Customer satisfaction. *Journal of Digitainability Realism & Mastery, 1*(01), 1-11. https://doi.org/10.56982/journalo.v1i01.6

Eusebio, E. J., Faeldan, F., & Moreno, E. (2023). Emotional Intelligence, Consumer-Perceived Values, and E-Shopping Satisfaction: A Mediation analysis. *Romblon State University Research Journal*, *4*(2), 1-5. https://doi.org/10.58780/rsurj.v4i2.84

Goffman, E. (2017). *Interaction ritual: Essays in face-to-face behavior*. Routledge.

https://doi.org/10.4324/9780203788387

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*(4), 370-396. https://doi.org/10.1037/h0054346

Rousseau, D. (1995). *Psychological contracts in organizations: Understanding written and unwritten agreements*. Sage. https://doi.org/10.4135/9781452231594

Karthik, J. P., & Sreehari, S. (2025). A STUDY ON CONSUMER TRUST IN ONLINE AND OFFLINE COMMERCE. *EPRA International Journal of Environmental Economics, Commerce and Educational Management, 12*(5), 229-232. https://doi.org/10.36713/epra21894

Gelashvili, V., Martínez-Navalón, J. G., DeMatos, N., & de Brito Correia, M. (2024). Technological transformation: The importance of E-WOM and perceived privacy in the context of opinion platforms. *Technological Forecasting and Social Change, 205*, 123472. https://doi.org/10.1016/j.techfore.2024.123472

Ortega-Bolaños, R., Bernal-Salcedo, J., Germán Ortiz, M., Galeano Sarmiento, J., Ruz, G. A., & Tabares-Soto, R. (2024). Applying the ethics of AI: a systematic review of tools for developing and assessing AI-based systems. *Artificial Intelligence Review, 57*(5), 1-30. https://doi.org/10.1007/s10462-024-10740-3

Ali, N. S. M. S. (2024). Adoption of Self-Checkout systems in retail and their impact on customer experience. *Journal of Informatics Education and Research*, *4*(3). 3575-3581. https://doi.org/10.52783/jier.v4i3.1970

Chen, I.-C., & Lu, J. (2025). Exploring User Experience Perceptions of Logistics E-commerce Platforms from the Perspective of E-commerce Politeness. *International Journal of Social Science and Economics Invention, 11*(04)*,* 90-96. https://doi.org/10.23958/ijssei/vol11-i04/393

Larivière, B., Bowen, D., Andreassen, T. W., Kunz, W., Sirianni, N. J., Voss, C., Wünderlich, N. V., & De Keyser, A. (2017). "Service Encounter 2.0": An investigation into the roles of technology, employees and customers. *Journal of Business Research*, *79*, 238-246. https://doi.org/10.1016/j.jbusres.2017.03.008

Sahu, S. (2025). The Role of AI Chatbots in Enhancing Online Customer Experiences. *International Scientific Journal of Engineering and Management, 04*(06)*,* 1-9. https://doi.org/10.55041/isjem04280

Rasheed, M., Sami, I., & Tabassam, A. (2025). The Impact of AI-Powered Chatbots on Customer Satisfaction and Business Performance in E-Commerce. *Social Science Review Archives*, *3*(1), 2390-2401. https://doi.org/10.70670/sra.v3i1.547

Clelland, C. L., Moss, S., & Clelland, J. D. (2024). Warning: Artificial intelligence chatbots can generate inaccurate medical and scientific information and references. *Exploration of Digital Health Technologies, 2*(1), 1-6. https ://doi.org/10.37349/edht.2024.00006

Abiagom, C. N., & Ijomah, T. I. (2024). Enhancing customer experience through AI-driven language processing in service interactions. *Open access research journal of engineering and technology*, *7*(1), 014-021. https://doi.org/10.53022/oarjet.2024.7.1.0027

Xie, Y., Tong, Z., & Wu, Z. (2025). Artificial Intelligence or Human Service, Which Customer Service Failure Is More Unforgivable? A Counterfactual Thinking Perspective. *Psychology & Marketing, 42*(8), *1993-2011.* https://doi.org/10.1002/mar.22215

Zhao, L., & Wu, W. (2025). Is Customer Service Innovation Always Preferable? The Impact of AI Technology on an Online Retailer's Customer Service Decision. *Journal of Theoretical and Applied Electronic Commerce Research, 20(2)*, 101. https://doi.org/10.3390/jtaer20020101

Gremler, D. D. (2015). The critical incident technique. In C. L. Cooper (Ed.), *Wiley encyclopedia* *of management.* Wiley. https://doi.org/10.1002/9781118785317.weom090062

Jombe, M. K., & Pretorius, M. (2025). Exploring the critical incident technique to inform distress causality and sensemaking. *Acta Commercii-Independent Research Journal in the Management Sciences*, *25*(1),1-12. https://hdl.handle.net/10520/ejc-acom_v25_n1_a1312

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin, 51*(4), 327. https ://doi.org/10.1037/h0061470

Rotem, S.-H., Potari, D., & Psycharis, G. (2024). Using critical incidents as a tool for promoting prospective teachers' noticing during reflective discussions in a fieldwork-based university course. *Educational Studies in Mathematics, 117*(1), 67-95. https://doi.org/10.1007/s10649-024-10336-2

Baker, M. A., & Kim, K. (2019). Value destruction in exaggerated online reviews. *International Journal of Contemporary Hospitality Management, 31*(4), 1956-1976. https://doi.org/10.1108/ijchm-03-2018-0247

Durosini, I., Triberti, S., Savioni, L., & Pravettoni, G. (2021). In the eye of a quiet storm: A critical incident study on the quarantine experience during the coronavirus pandemic. *PloS one*, *16*(2), 1-17. https://doi.org/10.1371/journal.pone.0247121

Madsen, H. L., & Ulhøi, J. P. (2021). Sustainable visioning: Re-framing strategic vision to enable a sustainable corporate transformation. *Journal of Cleaner Production*, *288*, 125602. https://doi.org/10.1016/j.jclepro.2020.125602

Bitner, M. J., Booms, B. H., & Tetreault, M. S. (1990). The Service Encounter: Diagnosing Favorable and Unfavorable Incidents. *Journal of Marketing*, *54*(1), 71-84. https://doi.org/10.1177/002224299005400105

Artino, A. R., La Rochelle, J. S., Dezee, K. J., & Gehlbach, H. (2014). Developing questionnaires for educational research: AMEE Guide No. 87. *Medical Teacher*, *36*(6), 463-474. https://doi.org/10.3109/0142159X.2014.889814

Ansari, M. M., & Khan, S. (2023). An in-depth examination of validity assessment: Exploring diverse methodologies and dimensions of validity in social research studies. *Asian Journal of Agricultural Extension, Economics & Sociology*, *41*(10), 772-782. https://doi.org/10.9734/ajaees/2023/v41i102224

Hirano, H., & Roberson, C. (2010). Measures, validity, and reliability. In H. Hirano & C. Roberson (Eds.), *Criminal justice research methods* (pp. 97-110). Routledge. https://doi.org/10.4324/9781439894880-7

Lim, W. M. (2024). A typology of validity: content, face, convergent, discriminant, nomological and predictive validity. *Journal of Trade Science*, *12*(3), 155-179. https://doi.org/10.1108/JTS-03-2024-0016

Olanipekun, A., Ahmed, V., Opoku, A., & Sutrisna, M. (2022). Understanding validity in research. In *Validity and Reliability in Built Environment Research: A Selection of Case Studies* (pp. 16-26). CRC Press. https://doi.org/10.1201/9780429243226-3