Electronic Markets (2026) 36:62
https://doi.org/10.1007/s12525-026-00918-8

RESEARCH PAPER

Can AI chatbots recover service failures? Exploring user experience and continuance intentions

Saule Yolcu[1] · Alperen Şahin[2] · Aslı Erdoğdu[3] · Taşkın Dirsehan[4,5,6]

Received: 22 November 2025 / Accepted: 24 June 2026

Abstract
Mobile food ordering applications (MFOAs) increasingly use AI-powered chatbots for customer interaction, making the quality of the user experience crucial for adoption and continued use. This study explores how user experience (UX) with MFOA chatbots influences satisfaction and continuance intention. It aims to reveal how both pragmatic (functional) and hedonic (emotional) UX dimensions shape ongoing user engagement with these AI-driven service technologies. A quantitative online survey was administered via Qualtrics between April and May 2025, targeting MFOA chatbot users in Turkey. Participants were recruited using convenience sampling and were screened to ensure prior experience with chatbot-based complaint handling in MFOAs. Data from 227 valid responses were analyzed using PLS-SEM. The results indicate that user experience, comprising both pragmatic and hedonic dimensions, significantly and positively affects satisfaction. Satisfaction fully mediates the relationship between user experience and continuance intention, with no significant direct effect of UX on continuance intention observed, thereby confirming satisfaction's central role in sustaining post-adoption behavior. This research contributes empirical evidence by integrating UX theory with post-adoption models in the context of AI-enabled service interactions. The findings emphasize the practical necessity for managers and developers to design chatbot experiences that balance functionality and enjoyment, thereby strengthening customer engagement and loyalty in mobile service environments.

Keywords Complaint handling · User experience · Customer satisfaction · Service recovery · Mobile food ordering apps

JEL Classification M30 · L81 · L84

Introduction

Effective customer service has emerged as a strategic necessity for businesses looking to improve customer happiness and cultivate enduring loyalty in the modern digital economy. It is often acknowledged that a major factor in determining both customer retention and brand perception is the quality of services provided (Wang et al., 2025). Businesses are finding it more difficult to handle an increasing number of consumer interactions while preserving responsiveness, accuracy, and

Responsible Editor: Yunkai Zhai

✉ Taşkın Dirsehan
taskin.dirsehan@marmara.edu.tr; dirsehan@essb.eur.nl

Saule Yolcu
s.momynova@almau.edu.kz

Alperen Şahin
alperensahin@esenyurt.edu.tr

Aslı Erdoğdu
asli.erdogdu@kent.edu.tr

[1] School of Management, Almaty Management University, Almaty, Kazakhstan

[2] Faculty of Business and Management Sciences, İstanbul Esenyurt University, Istanbul, Turkey

[3] Vocational School of Health Services, Istanbul Kent University, Istanbul, Turkey

[4] Al-Farabi Kazakh National University, Almaty, Kazakhstan

[5] Faculty of Business Administration, Marmara University, Göztepe Campus, Kadıköy, Istanbul 34722, Turkey

[6] Erasmus School of Social and Behavioural Sciences, Rotterdam, the Netherlands

Published online: 08 July 2026

personalization as digital communication channels multiply. Even while conversational AI technologies present exciting opportunities for automation, conventional systems frequently fall short of the intricate and subtle requirements of human language, especially given the high standards of modern customers (Fürst et al., 2025; Kim, 2023; Lee, 2024). These difficulties highlight the necessity of more sophisticated, user-centered methods for designing and evaluating chatbots in customer support settings (Shareef, 2024).

AI-powered chatbots are one of the major technological advancements influencing the consumer experience (Aslam, 2023; Bhattaru et al., 2024). According to recent survey findings, consumers have increasingly embraced automated customer care services. A survey conducted by Tidio and Statista found that 96% of respondents believe that more organizations should adopt chatbots in place of traditional customer service agents. Additionally, 94% expect chatbot technology to eventually replace conventional contact centers, while 82% prefer resolving their issues through chatbots rather than waiting to speak with a customer support representative. These findings highlight the growing acceptance of AI-powered customer service solutions and the changing expectations of modern consumers (Statista, 2025).

Chatbots are not limited to ordinary activities, but also assist with service failure recovery processes (Lee, 2024). Chatbots' growing popularity in service encounters can be ascribed to their efficiency, availability, and capacity to solve problems in real time. Furthermore, the design architecture of advanced conversational agents, such as Microsoft's Xiaoice, prioritizes both technical competence and emotional intelligence. Chatbots may now give more contextually appropriate, inspiring, and positive responses thanks to the incorporation of empathy and social capabilities such as user profiling, sentiment detection, and emotion analysis. Even service-oriented or transactional bots have begun to create their own personalities to improve the user experience (Mehra, 2021). This capacity is crucial in the mobile food delivery industry, which frequently experiences service problems such as delayed orders or incorrect deliveries. User expectations of food delivery apps are changing rapidly due to variables such as 24/7 working hours, urbanization, rising disposable income, more working women, and widespread smartphone use. They require quickness, convenience of ordering in advance, transparent menu details, promotional offers, loyalty reward systems, and grievance redressal processes (Chandramouli, 2019). The global online meal delivery service has seen fast expansion, with revenue estimated to reach US$1.65 trillion by 2027 (Lefebvre et al., 2024).

MFOAs are regularly upgraded in terms of service, functionality, design, and content to keep up with this expansion (Ghali, 2025). In this dynamic setting, innovative chatbot systems that manage service recovery procedures and enhance conversation quality might provide a major theoretical linkage by impacting perceived fairness, customer happiness, and loyalty intentions (Park et al., 2025). Although internet delivery services, notably MFOAs, have attracted substantial attention in Turkey and other countries in the area, scholars and researchers have not adequately investigated the elements impacting users' views, intentions, and behaviors toward these apps (Dirsehan & Cankat, 2021). Furthermore, previous research has primarily focused on customers' first approval or adoption of the technology (Yoon & Yu, 2022). To reap the projected benefits of chatbots in customer service, users must have favorable interactions with these technologies. This is especially crucial because the adoption rate of customer support chatbots among users is still below industry expectations (Arif et al., 2024; Følstad & Taylor, 2021). This circumstance mandates that service providers improve their capability to measure and monitor user experience. This allows for a more effective identification of key elements influencing user experience, the support of continuous improvement efforts, and a better knowledge of user behaviors and their changes over time (Kvale et al., 2020).

While past research on chatbots has mostly focused on adoption, usability, and overall satisfaction, little is known about how chatbot-based complaint handling in service failure scenarios differs from traditional service recovery theories. To address this gap, the current study develops and tests a conceptual model that links the hedonic and pragmatic dimensions of user experience to satisfaction and continuance intention in the context of AI-mediated complaint management. By focusing on consumers who use chatbots to resolve service-related issues, it examines how post-failure interactions are evaluated and how these evaluations influence both satisfaction and continuance intention.

As a result, this study provides new insights into the mechanisms of chatbot-mediated complaint resolution and its role in enhancing long-term user engagement. Furthermore, by situating the analysis within the mobile food delivery industry, particularly in emerging markets like Turkey, the study extends the empirical scope of expectation-confirmation theory (ECT) and service recovery theory to AI-powered customer service interactions.

Accordingly, the study seeks to address the following research questions:

RQ1. How do the hedonic and pragmatic dimensions of user experience influence user satisfaction with chatbot-mediated complaint handling in MFOAs?
RQ2. To what extent does user satisfaction affect the continuance intention to use chatbots after service failures in MFOAs?
RQ3. Does user satisfaction mediate the relationship between user experience (hedonic and pragmatic) and continuance intention in AI-driven complaint handling contexts?

To guide the investigation of these questions, the study integrates Service Recovery Theory and Expectation-Confirmation Theory (ECT) as its theoretical foundation, providing a comprehensive lens through which user experiences with chatbots are linked to satisfaction and continuance intentions in post-failure contexts. The findings of this study carry important theoretical implications. By demonstrating that both hedonic and pragmatic UX dimensions influence satisfaction, which in turn drives continuance intention, this study extends expectation-confirmation theory to AI-mediated complaint handling contexts. Furthermore, the confirmation of satisfaction as a full mediator broadens the applicability of service recovery theory to chatbot-based interactions, suggesting that recovery quality perceptions operate similarly in human and AI-driven service encounters.

From a practical standpoint, the results suggest that MFOA service providers should prioritize both functional efficiency and emotional engagement when designing chatbot interactions. Specifically, managers and developers should invest in chatbot interfaces that not only resolve complaints quickly and accurately but also create enjoyable and engaging experiences. Given that satisfaction fully mediates the path to continued use, ensuring positive post-failure experiences is critical for retaining users and sustaining long-term engagement in mobile food delivery markets.

Literature review and hypothesis development

In line with PRISMA guidelines (Page et al., 2021), a systematic literature review was conducted using the Scopus and Web of Science databases on March 7, 2025. Studies containing the terms "chatbot" and "user experience" in their titles, abstracts, or keywords were identified. The initial search yielded 1,042 records (696 from Scopus, 346 from Web of Science). After applying language and publication type filters, 407 studies remained (238 from Scopus, 169 from Web of Science). These were merged and deduplicated using bibliometric tools (Kara et al., 2025), resulting in 257 unique records. Through full-text screening (Şahin et al., 2025) based on relevance, 130 studies were included in the final review.

However, while these 130 studies generally focus on user experience in chatbots, none of them specifically address user experience in MFOAs, which is the focus of this study. Only 13 studies discussed topics related to customer service, which are closely aligned with the service failure recovery aspect explored in this study. These studies are presented in Table 1.

Table 1 primarily presents studies focusing on chatbots and user experience in customer service, with an emphasis on user satisfaction, interaction quality, and the effectiveness of AI in service delivery. While the majority of the studies explore general aspects of chatbot interactions, user perceptions, and technological integrations, only a few directly address service failures.

Although the search strategy employed broad keywords related to chatbot user experience, this was a deliberate methodological choice consistent with PRISMA-based systematic review practices. Given the scarcity of studies explicitly examining chatbot-mediated complaint handling in MFOAs, an initially inclusive search strategy was used to avoid the exclusion of conceptually relevant work. The retrieved studies were subsequently subjected to careful manual screening, and only those theoretically aligned with customer service, service failure, and post-failure user responses were retained. Insights from this broader literature were then theoretically mapped onto the MFOA context by focusing on complaint handling, service recovery interactions, and continuance-related outcomes specific to mobile food ordering platforms. For instance, findings on personalization and interaction quality in delivery platforms (Mustafa, 2024) and emotional coping in chatbot service failures (Zhang et al., 2024) were directly mapped onto the complaint-handling dynamics specific to MFOA settings. This approach ensured comprehensive coverage while maintaining contextual relevance.

Role of chatbots in handling service failures

Service failure is described as a service performance that falls short of the customer's expectations or the acceptable quality of service. A mismatch between expectations and actual service experience can lead to client discontent, which may result in undesirable outcomes such as decreased customer loyalty, customer attrition, and unfavorable word-of-mouth communication (Rese & Witthohn, 2025). Service failures are often divided into two categories: process failures and result failures (Choi et al., 2021; Sands et al., 2022). Process failures are disruptions or inadequacies in the delivery of the service, whereas outcome failures occur when the service's fundamental needs are not met. Process failures are frequently associated with the attitude or conduct of AI chatbot services, whereas outcome failures are linked to the chatbot's functional capabilities (Choi et al., 2021).

Customer complaints are the behavioral responses of unsatisfied customers to service failures. When a consumer is dissatisfied with a service failure, they frequently blame the guilty party. Furthermore, service failures are extremely likely to result in consumer discontent and negative word-of-mouth, which can harm a brand's reputation and trigger customer loss. As a result, firms strive to respond quickly to service problems. Service recovery reduces the negative impact of service failures on customers, develops customer relationships, and increases trust in the firm (Lee, 2024).

In recent years, advancements in artificial intelligence and natural language processing have permitted the introduction

Table 1 Customer service-related studies focusing on service failure recovery in the context of chatbots and user experience

Authors	Methods/approach	Key contribution or perspective
(Følstad & Taylor, 2021)	Longitudinal collaborative analysis	Synthesizes long-term academic discussions to outline future research directions in chatbot and dialogue system studies
(Bhattacharya & Sinha, 2022)	Interviews, quantitative survey	Analyzes how AI-powered chatbots influence customer experience and satisfaction in digital banking services
(Diederich et al., 2022)	Thematic clustering analysis	Provides an organizing review of chatbot and virtual human studies, mapping conceptual themes in the field
(Haugeland et al., 2022)	Randomized experiment, semi-structured interviews	Investigates how human-like features and social presence affect user experience in chatbot-based interactions
(Ltifi, 2023)	Survey, SEM (structural equation modeling)	Tests moderating effects of empathy, usability, and trust on user experience with chatbots
(Mustafa, 2024)	Qualitative and quantitative surveys	Investigates the role of personalization and AI-driven interaction quality in enhancing user experience on delivery platforms
(Rajabi et al., 2024)	Knowledge graph analysis	Reviews AI-based virtual assistants using knowledge graph frameworks
(Saklani & Kala, 2024)	Semi-structured interviews, thematic analysis	Explores generational attitudes and trust toward chatbots among Gen Z consumers in the Indian context
(Sohail et al., 2024)	Field survey	Tests a model integrating ECT and TTF to explain satisfaction and continuance intentions in intelligent agent use
(Wut et al., 2024)	Survey	Examines how chatbot-based customer service interactions influence satisfaction and electronic word-of-mouth intentions
(Zhang et al., 2024)	Thematic content analysis	Explores emotional coping mechanisms and user reactions to service failures in chatbot interactions
(Meng et al., 2025)	Scenario-based experiment	Examines how chatbot response strategies and emotional cues influence purchase intention
(Wang & Lo, 2025)	2 × 2 experimental design	Investigates how delayed chatbot responses affect satisfaction across age groups

of powerful chatbots with human-like conversational abilities and personalized help. Systems like ChatGPT, also known as large language models, have revolutionized conversational AI by training on large datasets and obtaining the ability to generate human-like responses by comprehending context (Kang & Hong, 2025). Chatbots are AI-powered systems that reply to user requests via text-based interactions. However, these systems do not always work smoothly, and service outages are very regular (Xing, 2024).

Chatbots are AI systems that process, analyze, and respond intelligently to natural language. This enables users to speak with intelligent systems in natural conversational languages, as if they were conversing with a human (Stefanidi et al., 2019). Innovations in artificial intelligence technologies, particularly machine learning, natural language processing (NLP), and intelligent analytics, have been critical milestones in chatbot development. Thanks to these developments, chatbots can learn continuously and leverage prior conversational experiences to generate more relevant and context-aware responses (Rahmanti et al., 2022). These technological and behavioral dynamics highlight the significance of optimizing chatbot interactions, especially in service failure recovery circumstances (Ghali, 2025).

The usage of chatbots in food ordering and complaint management is especially important when dealing with service failures, which can impact consumer satisfaction and brand loyalty. Recent research indicates that enhanced interaction quality in chatbots has a considerable impact on consumers' views of fairness, contentment, and loyalty (Park et al., 2025). Such shortcomings in chatbot performance might impact user happiness and the desire to reuse the chatbot. Poor service delivery frequently results in chatbots being viewed as less "human-like," leading to lower levels of trust and consumer loyalty (Rese & Witthohn, 2025). Empathy-based messages are excellent in establishing emotional relationships, whereas solution-oriented solutions are better at generating technical trust. The frequency and reason of the failure frequently influence which method should be employed. For example, customers typically anticipate practical answers in the case of repeated failures, whereas empathy is more appropriate in isolated incidents (Haupt et al., 2023).

AI chatbots are particularly effective in compensating for human-caused service failures, especially when they provide fair and prompt responses that increase user satisfaction. Furthermore, the ability of AI to interact empathically and effectively improves the success of recovery attempts (Lee, 2024). Chatbots' function in service failure recovery has been connected to customer loyalty using the frustration-aggression theory. Effective and consistent responses from chatbots have been demonstrated to improve loyalty, but insufficient interactions tend to intensify negative emotions and impair commitment (Ozuem et al., 2024). Non-functional chatbot failures, such as communication mistakes, have been shown to have a greater impact on recovery expectations than technical faults. Users' views of controllability, the severity of the failure, and the chatbot's human-like features all influence this effect (Song et al., 2023).

In contrast, little study has been conducted on chatbots' ability to recover from service outages. Service recovery is a critical activity required to rectify flaws in the service delivery process and transform service failure into positive outcomes. Creating a successful service recovery strategy is crucial for retaining customers and establishing good connections. Genuine recovery, such as customer forgiveness and post-recovery satisfaction, is typically the ultimate goal of customer service efforts aimed at restoring customer relationships after a service failure.

Understanding the dynamics of user experience in relation to customer expectations in chatbot use

The use of chatbots has been shown to enhance user satisfaction, particularly by improving system usability and information quality (Merkouris et al., 2024). These findings highlight a strong connection between chatbot interaction and user satisfaction (Chumkaew, 2023; Macedo et al., 2024). Core features such as perceived usefulness, ubiquitous access, completeness, accuracy and convenience have been shown to significantly enhance satisfaction (Alotaibi & Hidayat-ur-Rehman, 2025; Chagas et al., 2023). user experience has also been significantly linked to performance expectancy, effort expectancy, trust, intention to use and chatbot usage (Dawar et al., 2022). These findings collectively support the conclusion that chatbot usage positively influences user experience and that users report high levels of satisfaction when interacting with such systems.

User satisfaction is closely associated with a positive user experience (Kim et al., 2025; Macedo et al., 2024; Wut et al., 2024). Dimensions of user experience such as usability, interactivity and efficiency have been found to significantly influence user satisfaction. While the rapid response capabilities of chatbots and their ability to resolve order-related issues tend to enhance satisfaction, challenges such as irrelevant replies, limited functionality, and the inability to handle complex queries may negatively impact the user experience. Notably, the use of advanced natural language processing (NLP) techniques and more human-like interactions has been emphasized as critical for improving user satisfaction. These findings further reinforce the positive effect of user experience on user satisfaction (Mustafa, 2024). Furthermore, faster response times and lower error rates contribute to higher levels of user satisfaction. Collectively, this evidence highlights the meaningful effect of user experience on user satisfaction and underscores the potential of effective UX design in enhancing overall satisfaction (Alabbas & Alomar, 2025; Entenberg et al., 2023; Kang & Hong, 2025; Mejia et al., 2024; Pezenka et al., 2024).

Theoretically, this study builds on two complementary perspectives. First, service recovery theory (Smith et al., 1999; Tax & Brown, 1998) suggests that customers' perceptions of fairness and resolution shape post-failure satisfaction and loyalty. Second, expectation-confirmation theory (ECT) (Bhattacherjee, 2001; Oliver, 1980) posits that satisfaction emerges when perceived performance meets or exceeds expectations. By applying these perspectives to chatbot-mediated interactions, this study examines whether user experience translates into continuance intention directly or only through satisfaction.

Service recovery theory provides a foundation for understanding how users evaluate service encounters following a failure, particularly in terms of perceived responsiveness, fairness, and problem resolution. In chatbot-based complaint handling, these evaluations are not limited to outcomes but are shaped by the experiential qualities of the interaction. In this regard, hedonic experiences capture the affective and emotional responses elicited by the chatbot, while pragmatic experiences reflect users' assessments of task efficiency and problem-solving effectiveness during recovery.

Expectation-confirmation theory complements this perspective by explaining how these experiential evaluations are cognitively processed and translated into satisfaction and continuance intention. Specifically, users compare their post-recovery experiences with prior expectations, and satisfaction emerges from the confirmation or disconfirmation of these expectations, subsequently influencing continuance intention. Accordingly, hedonic and pragmatic experiences operate as experiential mechanisms linking service recovery perceptions to expectation confirmation and post-recovery behavioral intentions.

Because service recovery theory focuses primarily on post-failure evaluations and expectation-confirmation theory explains post-evaluative satisfaction and behavioral outcomes, neither framework alone sufficiently captures the full experiential–evaluative–behavioral process underlying chatbot-mediated complaint handling. Their integration therefore offers a more comprehensive theoretical explanation of how users experience, evaluate, and decide to continue using AI-based complaint handling systems.

This theoretical integration refines existing models by highlighting the mediating role of satisfaction in digital complaint resolution, particularly in high-failure environments such as food delivery services. Based on these insights, the following hypothesis has been formulated:

H1: MFOA chatbot user experience positively affects MFOA chatbot satisfaction.

Various factors such as perceived usefulness, user satisfaction, completeness of information and accessibility shape users' intentions to use chatbots. In particular, perceived usefulness and user satisfaction have been found to positively influence users' intentions to engage with chatbots. Moreover, user satisfaction itself has a direct and significant impact on the intention to use chatbots (Alotaibi & Hidayat-ur-Rehman, 2025). The capabilities of chatbots, the quality of information they provide and their ability to simulate human-like interactions contribute positively to the overall user experience. These aspects, in turn, enhance users' intentions to continue using the chatbot (Kim, 2024; Wut et al., 2024). Customer loyalty also exerts a positive effect on continued usage intention (Cheng & Jiang, 2020; Shahzad et al., 2024). These findings collectively support the notion that user experience has a positive and significant influence on the intention to continue using chatbots (Wut et al., 2024). Based on these insights, the following hypotheses have been developed:

H2: MFOA chatbot satisfaction positively affects MFOA chatbot continuance intention.
H3: MFOA chatbot user experience positively affects MFOA chatbot continuance intention.
H4: MFOA chatbot satisfaction mediates the relationship between user experience and continuance intention.

The research model comprises three primary constructs: user experience, satisfaction, and continuance intention. These constructs reflect the sequential process by which users evaluate interactions with the chatbot and form behavioral intentions. Based on prior literature, three hypotheses were formulated to test the relationships among these constructs (see Fig. 1).

As illustrated in Fig. 1, the conceptual model examines the effects of hedonic and pragmatic experiences on user satisfaction and continuance intention. All constructs were measured reflectively, with latent variables represented by multiple observed indicators. Pragmatic experience was measured through UE1–UE4, hedonic experience through UE5–UE8, satisfaction via CS1–CS3, and continuance intention via CI1–CI4. In this approach, changes in latent variables are assumed to drive changes in the associated indicators, making it suitable for capturing user perceptions and behavioral tendencies.

Methods

This study employed a quantitative research design to investigate the impact of user experience on satisfaction and continuance intention in the context of MFOA chatbots. Data were collected via a structured online questionnaire consisting of three sections. The first section included screening questions to ensure participant eligibility, specifically assessing prior experience with the MFOA chatbot and their use of it for handling complaints. The second section focused on evaluating users' experiences with the chatbot, forming the basis for testing the proposed research model. The final section gathered information on behavioral intentions, demographic characteristics (age, gender), and socioeconomic status, including education, occupational status, monthly expenditure, and city of residence.

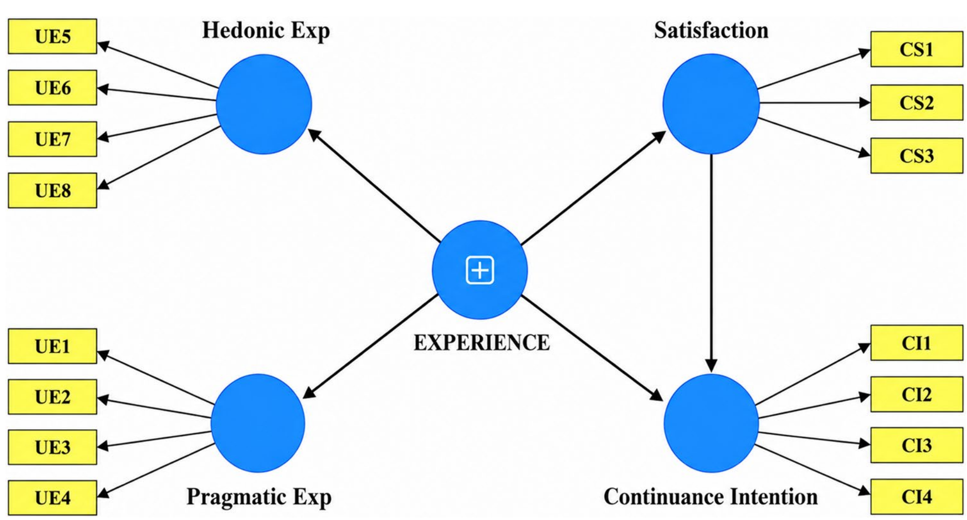

Fig. 1 Conceptual model

Measurement development

Constructs were measured using established scales to ensure content validity. User experience was assessed with eight items adapted from the User Experience Questionnaire (UEQ-S) (Hinderks & Thomaschewski, 2017), using a five-point semantic differential scale (e.g., obstructive–supportive, complicated–easy, boring–exciting). Satisfaction was measured with three items adapted from (Hsiao et al., 2016), and continuance intention with four items adapted from (Melián-González et al., 2021). Although these scales were originally developed in online shopping and hotel chatbot contexts respectively, all items were adapted to reflect the MFOA chatbot complaint-handling context. Prior to the main survey items, participants were explicitly instructed to evaluate all subsequent questions based on their experience with the chatbot service used in their most recent mobile food order. This contextual framing ensured that respondents interpreted all scale items within the intended MFOA complaint-handling setting, regardless of the original context in which the scales were developed. A pretest with 15 participants was conducted, and minor adjustments were made to improve clarity.

Translation and adaptation procedure

Because the study targeted Turkish participants, a double-translation procedure was applied to ensure linguistic and conceptual equivalence (Kristjansson et al., 2003; McGorry, 2000). Two bilingual academics (one in marketing communication, one in information systems) independently translated the original English items into Turkish. Two additional bilingual experts then back-translated the items into English. Discrepancies were discussed in a joint session, and minor wording adjustments were made for cultural appropriateness and clarity (e.g., replacing "online shopping" with "food delivery via mobile apps").

Data collection and sample characteristics

The survey was administered online via Qualtrics, targeting users familiar with MFOA chatbots. As all participants were chatbot users, digital literacy concerns were considered minimal (Malhotra, 2020). Participants were drawn from multiple MFOA platforms. While these platforms may employ different chatbot architectures or underlying LLMs, the study focused on users' perceived experience outcomes, rather than technical characteristics of the chatbot systems. The final sample comprised 227 participants (64.8% female, 35.2% male), primarily aged 18–24 (53.3%) and 25–34 (28.6%). Educational backgrounds included associate degree students (30.4%), bachelor's degree holders (11.9%), and master's students (11.9%). Monthly expenditures were ≤ 5,000 TL for 57.3% of respondents. The most frequently used mobile platforms for food delivery were Trendyol (n = 41), Yemeksepeti (n = 23), and Getir (n = 13). Recent engagement was high, with 57.2% having ordered within the past week.

To ensure participants had actual experience with chatbot-based complaint handling in MFOAs, we implemented a multi-stage screening process. First, participants were asked: "Do you use chatbots to resolve issues you encounter with your mobile food orders?" Only those who responded affirmatively proceeded to the survey. Second, participants indicated the types of complaints they had encountered and attempted to resolve through chatbots (e.g., delayed delivery, incorrect orders, food quality issues). Third, we measured application usage recency through the question "When did you last use the application?" to ensure participants were active users with recent MFOA experience. Only participants who confirmed chatbot usage for complaint handling, could identify specific complaint types, and demonstrated recent application usage were included in the final sample (N = 227).

Partial least squares structural equation modeling (PLS-SEM) was employed using SmartPLS 4 (Ringle et al., 2024). Following established guidelines, a minimum of 200 respondents were considered sufficient (Hair et al., 2011; Hoelter, 1983). The study's sample of 227 participants exceeded this threshold, ensuring adequate statistical power. Additionally, both the 10-times rule (Hair et al., 2011) and the minimum R^2 approach (Cohen, 2013; Sarstedt et al., 2014) confirmed that the sample size was sufficient for the proposed model.

As shown in Table 2, pragmatic and hedonic user experience items capture users' task-oriented and pleasure-oriented perceptions, respectively. Satisfaction items reflect overall evaluations of the chatbot service, while continuance intention items indicate participants' likelihood of future use. The use of validated scales and reflective measurement ensures content validity and reliability, providing a robust foundation for subsequent PLS-SEM analyses.

The robust data collection and well-defined measurement model established a solid foundation for testing the proposed hypotheses. With a sufficient sample size and validated reflective constructs, the dataset was suitable for partial least squares structural equation modeling (PLS-SEM) analyses. The following section presents the results of these analyses, including assessments of the measurement model, structural relationships among constructs, and the empirical testing of the proposed hypotheses. These findings provide insights into the influence of hedonic and pragmatic user experiences on satisfaction and continuance intention in the context of MFOA chatbots.

Table 2 List of constructs and their items

Construct	Items	References
Pragmatic user experience Hedonic user experience	*MFOA_UES_1*. obstructive-supportive	adapted from (Hinderks & Thomaschewski, 2017)
	MFOA_UES_2. complicated-easy	
	MFOA_UES_3. inefficient-efficient	
	MFOA_UES_4. confusing-clear	
	MFOA_UES_5. boring-exciting	
	MFOA_UES_6. not interesting-interesting	
	MFOA_UES_7. conventional-inventive	
	MFOA_UES_8. usual-leading edge	
Satisfaction	*MFOA_CS_1*. I believe that using a chatbot in my online shopping was a right decision	adapted from (Hsiao et al., 2016)
	MFOA_CS_2. My experience with using the chatbot was satisfactory	
	MFOA_CS_3. I am satisfied with the chatbot service I used	
Continuance intention	*MFOA_CI_1*. I intend to continue using chatbots in the future	adapted from (Melián-González et al., 2021)
	MFOA_CI_2. I will use chatbots when necessary	
	MFOA_CI_3. I plan to use chatbots in the future	
	MFOA_CI_4. I believe more and more people will use chatbots over time	

Note: Items in the table reflect the original English source scales. All items were translated into Turkish and adapted to the MFOA chatbot complaint-handling context following a double-translation procedure (see Sect. "Translation and adaptation procedure")

Construct reliability and validity

Before assessing the structural model, the measurement model was evaluated for reliability and validity. Internal consistency was examined using Cronbach's alpha, with values above 0.60 considered satisfactory (Nunnally, 1978). Convergent validity was assessed based on two criteria recommended by (Fornell & Larcker, 1981); (1) all factor loadings should be significant and exceed 0.70, and (2) the average variance extracted (AVE) for each construct should exceed 0.50. Discriminant validity was evaluated using the heterotrait-monotrait (HTMT) ratio, with a cutoff value of 0.85 (Henseler et al., 2015; Voorhees et al., 2016).

Table 3 presents the results of reliability and convergent validity analyses. Composite reliability (CR) and Cronbach's alpha coefficients for all constructs exceeded 0.70, demonstrating satisfactory internal consistency. Factor loadings ranged from 0.705 to 0.890, indicating acceptable indicator convergence. AVE values ranged from 0.507 to 0.715, exceeding the critical threshold of 0.50, as recommended by (Sarstedt et al., 2021), confirming that the majority of variance is captured by the latent constructs.

To assess construct reliability and validity, several statistical criteria were examined, including composite reliability (CR) for internal consistency, Cronbach's alpha, average variance extracted (AVE) for convergent validity, and outer loadings for indicator reliability.

The composite reliability (CR) values for all constructs were above the recommended threshold of 0.70 (Hair et al., 2019), indicating satisfactory internal consistency. Specifically, CR values were continuance intention (0.899), experience (0.890), hedonic experience (0.882), pragmatic experience (0.890), and satisfaction (0.883). Similarly, Cronbach's alpha values also exceeded 0.70 for all constructs, supporting the reliability of the measurement model. Convergent validity was confirmed through the average variance extracted (AVE), with all values surpassing the acceptable limit of 0.50. The AVE values were as follows: continuance intention (0.692), experience (0.507), hedonic experience (0.652), pragmatic experience (0.669), and satisfaction (0.715). These results indicate that a substantial portion of the variance is explained by the latent constructs. The outer loadings of the indicators were examined, and all items demonstrated loadings above the recommended threshold of 0.70, providing additional support for convergent validity. For example, the indicators of Continuance Intention ranged from 0.705 to 0.890, while those of Satisfaction ranged from 0.817 to 0.878.

Discriminant validity was further assessed using the heterotrait–monotrait ratio (HTMT), as recommended by Henseler et al. (2015). The HTMT values and their bootstrap confidence intervals are reported in Table 4.

As shown in Table 4, the HTMT values for all conceptually distinct construct pairs are below the recommended threshold of 0.90, and their bias-corrected bootstrap confidence intervals do not include the value of 1, indicating satisfactory discriminant validity. Higher HTMT values and confidence intervals approaching or slightly exceeding 1 are observed between hedonic experience, pragmatic experience, and overall experience. This result is theoretically

Table 3 Evaluation of scales for reliability and convergent validity

Construct	Original sample	Sample mean	Std. dev	*T* statistics	*p*-values	Cronbach's α	CR	AVE
CI1 ← continuance intention	0.886	0.882	0.049	18.033	0.000	0.899	0.899	0.692
CI2 ← continuance intention	0.705	0.703	0.068	10.302	0.000			
CI3 ← continuance intention	0.890	0.889	0.042	20.967	0.000			
CI4 ← continuance intention	0.833	0.836	0.040	13.799	0.000			
CS1 ← satisfaction	0.817	0.815	0.054	15.254	0.000	0.882	0.883	0.715
CS2 ← satisfaction	0.840	0.840	0.055	22.247	0.000			
CS3 ← satisfaction	0.878	0.878	0.049	17.940	0.000			
UE1 ← user experience (pragmatic)	0.843	0.843	0.042	18.903	0.000	0.890	0.890	0.507
UE2 ← user experience (pragmatic)	0.789	0.787	0.046	17.042	0.000			
UE3 ← user experience (pragmatic)	0.848	0.847	0.041	19.861	0.000			
UE4 ← user experience (pragmatic)	0.791	0.790	0.037	21.276	0.000			
UE5 ← user experience (hedonic)	0.881	0.881	0.057	13.158	0.000			
UE6 ← user experience (hedonic)	0.820	0.819	0.052	15.796	0.000			
UE7 ← user experience (hedonic)	0.813	0.812	0.045	18.057	0.000			
UE8 ← user experience (hedonic)	0.708	0.705	0.048	14.811	0.000			

Table 4 Discriminant validity assessment– heterotrait-monotrait ratio (HTMT)

	Original sample (O)	**Sample mean (M)**	**Bias**	**2.5%**	**97.5%**
EXPERIENCE ↔ continuance intention	0.293	0.310	0.017	0.177	0.430
Hedonic exp ↔ continuance intention	0.113	0.141	0.028	0.051	0.191
Hedonic exp ↔ EXPERIENCE	0.982	0.982	0.001	0.953	1.011
Pragmatic exp ↔ continuance intention	0.398	0.399	0.001	0.239	0.550
Pragmatic exp ↔ EXPERIENCE	0.980	0.981	0.001	0.951	1.008
Pragmatic exp ↔ hedonic Exp	0.586	0.585	0.000	0.447	0.702
Satisfaction ↔ continuance intention	0.712	0.712	0.001	0.582	0.810
Satisfaction ↔ EXPERIENCE	0.469	0.474	0.005	0.304	0.617
Satisfaction ↔ hedonic Exp	0.279	0.287	0.008	0.137	0.436
Satisfaction ↔ pragmatic Exp	0.538	0.539	0.001	0.370	0.678

expected, as hedonic and pragmatic experience represent first-order dimensions of the same higher-order experience construct. Therefore, the HTMT results among these sub-dimensions are not interpreted as evidence of insufficient discriminant validity but rather reflect the multidimensional nature of the experience construct. Notably, the HTMT value between hedonic and pragmatic experience (0.586) is well below the 0.85 threshold, providing empirical confirmation that these two constructs are distinct. Conceptually, hedonic experience captures affective and enjoyment-related aspects of chatbot interaction, while pragmatic experience reflects task-oriented and utilitarian dimensions, further supporting their discriminant validity.

The measurement model satisfies the required criteria for reliability, convergent validity, and discriminant validity. Accordingly, the results confirm the adequacy of the measurement model and support proceeding with the assessment of the structural model.

Common method bias and collinearity assessment

To assess potential common method bias (CMB), we employed two complementary approaches. First, we followed the full collinearity variance inflation factor (VIF) approach proposed by (Kock, 2015), which is specifically recommended for PLS-SEM studies as an effective alternative to confirmatory factor analysis-based methods. Construct-level collinearity was assessed using the inner model VIF values. The maximum VIF observed across all constructs was 1.31, well below the conservative threshold of 3.3, indicating that neither common method bias nor multicollinearity is likely to bias the results.

Second, we conducted Harman's single common factor test as an additional diagnostic (Podsakoff et al., 2003). All scale items were subjected to an exploratory factor analysis (principal axis factoring, no rotation) with the number of factors fixed at one. The single factor accounted for 37.8% of the total variance, well below the 50% threshold, providing further evidence against CMB.

Although we also attempted the unmeasured latent method factor (LMF) approach, this technique is primarily designed for covariance-based SEM (CB-SEM) and is not directly applicable to PLS-SEM contexts (Chin et al., 2012; Podsakoff et al., 2012). Consistent with this, our attempt to implement LMF via CFA yielded severe identification problems, with several factor loadings collapsing to near-zero or negative values — a known limitation of this technique in models with higher-order constructs and constructs with fewer than four indicators. Taken together, the full collinearity VIF and Harman's single factor test provide a robust and appropriate assessment of CMB for PLS-SEM-based research.

Results

Structural model assessment

The structural model was tested using partial least squares structural equation modeling (PLS-SEM) to estimate the structural relationships and test the hypotheses (Hair, 2014). PLS-SEM was chosen due to its suitability for analyzing complex models with hierarchical structures, particularly second-order formative constructs (Sarstedt et al., 2019), and its regression-based approach, which minimizes residual variances (Hair et al., 2011; Merli et al., 2018). PLS-SEM enables the simultaneous examination of direct, indirect, and mediating effects (Nitzl et al., 2016). To assess the explanatory power of the structural model, R^2 (coefficient of determination) values were examined. R^2 represents the proportion of variance in the endogenous constructs explained by their predictors. The results indicate that the R^2 values were 0.511 for continuance intention, 0.916 for hedonic experience, 0.999 for pragmatic experience, and 0.234 for satisfaction.

The very high R^2 values for hedonic experience (0.916) and pragmatic experience (0.999) should be interpreted in light of the hierarchical model specification. These constructs are modeled as first-order reflective dimensions that formatively define the second-order User Experience construct in a reflective–formative hierarchical component model (Type II) using the repeated indicators approach (Becker et al., 2012; Sarstedt et al., 2019). Accordingly, the R^2 values for these first-order constructs do not represent conventional predictive relationships but rather indicate the extent to which these experiential dimensions contribute to the higher-order User Experience construct. In formative second-order models, high sometimes near-perfect R^2 values for lower-order components are methodologically expected and theoretically appropriate, as the higher-order construct is constituted by its dimensions rather than predicting them in a causal sense (Hair, 2014). Pragmatic experience exhibits an almost complete contribution ($R^2 = 0.999$) to the higher-order construct. This pattern is theoretically plausible in complaint handling contexts, where users primarily evaluate chatbot interactions based on functional performance, including efficiency, clarity, ease of use, and task support. Hedonic experience ($R^2 = 0.916$), while still strongly represented, plays a complementary role by enhancing emotional engagement and perceived enjoyment rather than fully defining the overall experience (Van Vaerenbergh et al., 2019).

This distribution reflects the task-oriented nature of complaint handling, where utilitarian concerns dominate users' evaluations, while emotional engagement serves as an enhancing but secondary factor. Importantly, these high R^2 values do not indicate construct redundancy or multicollinearity. Variance inflation factor (VIF) values were well below critical thresholds (all < 1.31, Sect. "Common method bias and collinearity assessment"), and discriminant validity between hedonic and pragmatic experience was confirmed via the HTMT criterion (HTMT = 0.586, Table 4). These results demonstrate that the two dimensions remain empirically distinct despite their strong contribution to the higher-order User Experience construct. The model demonstrates adequate explanatory power for continuance intention ($R^2 = 0.511$) while providing a partial explanation of satisfaction ($R^2 = 0.234$), highlighting the central yet not exclusive role of experiential factors in chatbot-based complaint handling.

Path analysis and hypotheses testing

In this study, we tested several hypotheses related to the relationships between user experience, satisfaction, and continuance intention in the context of MFOA chatbot usage. The results of the path analysis are summarized in Table 5 and discussed below, along with the relevant path coefficients and statistical significance (Henseler et al., 2016; Sarstedt et al., 2014).

H1 posited that MFOA chatbot user experience leads to MFOA chatbot satisfaction. The path coefficient for this relationship was 0.483 ($p < 0.001$), which is statistically significant, supporting H1. This suggests that a positive user experience significantly enhances satisfaction with the MFOA chatbot.

H2 hypothesized that MFOA chatbot satisfaction leads to MFOA chatbot continuance intention. The path coefficient for this relationship was 0.742 ($p < 0.001$), which is also statistically significant, supporting H2. This indicates that higher satisfaction with the chatbot is a strong predictor of users' intention to continue using the chatbot.

H3 examined whether MFOA chatbot user experience leads directly to MFOA chatbot continuance intention. The path coefficient for this relationship was -0.060 ($p = 0.542$),

Table 5 Structural model standardized path coefficients (β)

Path	Path Coefficients	T Statistic	P Value
User Experience → Continuance Intention	-0.060	0.609	**0.542**
User Experience → Hedonic Exp	0.957	42.286	**0.000**
User Experience → Pragmatic Exp	0.999	82.217	**0.000**
User Experience → Satisfaction	0.483	5.719	**0.000**
Satisfaction → Continuance Intention	0.742	7.976	**0.000**

Note: Green paths indicate statistically significant path coefficients (p 0.05), whereas red paths indicate non-significant path coefficients ($p \geq 0.05$)

which is not statistically significant, thus failing to support **H3**. This suggests that user experience does not have a direct effect on continuance intention in this model.

H4 proposed that MFOA chatbot satisfaction mediates the relationship between user experience and continuance intention. While the direct effect of user experience on continuance intention was not significant ($\beta = -0.060$, $p = 0.542$), the significant effect of satisfaction on continuance intention ($\beta = 0.742$, $p < 0.001$) indicates a potential mediation pattern. This pattern is consistent with full mediation, which is formally tested in Sect. "Mediation analysis".

As hedonic and pragmatic experience are specified as first-order dimensions of the higher-order user experience construct, the very high path coefficients reflect their role as measurement dimensions rather than substantive causal effects. The findings underscore the pivotal role of satisfaction as a primary determinant of continuance intention, while also illustrating that user experience serves as a significant antecedent to satisfaction in the context of MFOA chatbot interactions.

Mediation analysis

To test the mediating role of satisfaction in the relationship between experience and continuance intention, we examined direct, indirect, and total effects with bootstrapped confidence intervals (5000 resamples) following the recommendations of (Preacher & Hayes, 2008).

As shown in Table 6, the results suggest a full mediation effect. The direct path from experience to continuance intention was not significant ($\beta = -0.060$, $p = 0.542$, 95% CI [−0.253, 0.138]), indicating no direct relationship. However, the indirect effect through satisfaction was positive and significant ($\beta = 0.359$, $p < 0.001$, 95% CI [0.208, 0.550]). The total effect was also significant ($\beta = 0.299$, $p = 0.001$, 95% CI [0.124, 0.471]).

These findings indicate that satisfaction fully mediates the relationship between experience and continuance intention. The bootstrapped confidence intervals for the indirect effect do not include zero, confirming the statistical significance of the mediation effect (Zhao et al., 2010). Accordingly, user experience influences continuance intention entirely through satisfaction rather than a direct path.

Our mediation analysis reveals a crucial finding: satisfaction fully mediates the relationship between user experience and continuance intention. The direct effect of experience on continuance intention was not significant ($\beta = -0.060$, $p = 0.542$), while the indirect effect through satisfaction was strong and significant ($\beta = 0.359$, $p < 0.001$). This full mediation pattern suggests that user experience with MFOA chatbots influences continuance intention entirely through its impact on satisfaction, rather than through a direct mechanism.

Table 6 Mediation analysis–effects of experience on continuance intention

Effect type	Path	β	t-value	p-value	95% CI	Result
Direct	EXPERIENCE → continuance intention	−0.06	0.609	0.542	[−0.253, 0.138]	Not supported
Indirect	EXPERIENCE → satisfaction → continuance intention	0.359***	4.13	< 0.001	[0.208, 0.550]	Supported
Total	EXPERIENCE → continuance intention	0.299**	3.326	0.001	[0.124, 0.471]	Supported

Note: Bootstrap samples = 5000. **$p < 0.01$; ***$p < 0.001$

Discussion

The findings show that user experience significantly contributes to satisfaction, although additional factors beyond UX also appear to influence satisfaction, demonstrating the importance of both hedonic and pragmatic dimensions in shaping post-failure evaluations. Hedonic experiences, defined by enjoyment, emotional engagement, and perceived social presence, improve the affective dimension of satisfaction, whereas pragmatic experiences, such as ease of use, efficiency, task completion, and responsiveness, improve the functional dimension of satisfaction. This dual-pathway influence emphasizes the importance of considering both the experience and utilitarian components of chatbots when building digital service interactions. Consistent with previous research (Chumkaew, 2023; Macedo et al., 2024; Merkouris et al., 2024), the study highlights that user experience is a central determinant of satisfaction in AI-mediated service contexts. These findings offer a more nuanced perspective of UX, stressing that emotional connection and practical functioning both influence how users perceive chatbot encounters, particularly in high-failure situations such as food delivery.

Most significantly, the study provides evidence that satisfaction fully mediates the relationship between UX and continuance intention. While previous research frequently assumed a clear relationship between great experience and continuous usage, our findings show that UX alone is insufficient to drive intention to continue using a chatbot-based service. The nonsignificant direct path from UX to continuance intention (−0.060) emphasizes the importance of satisfaction in converting positive experiences into behavioral intentions. This mediating role aligns with expectation-confirmation theory, which posits that satisfaction emerges when perceived performance meets or exceeds expectations (Fu et al., 2018). It also complements service recovery theory, which emphasizes the influence of perceived fairness and effective resolution on post-failure evaluations (Van Vaerenbergh et al., 2019). By empirically proving this mediation, the study extends theoretical knowledge by demonstrating that in AI-mediated complaint management, satisfaction serves as the primary channel connecting user experience to continuous involvement.

The strong path coefficients between UX and hedonic (0.957) and pragmatic (0.999) experiences further reinforce the importance of considering the multidimensional nature of UX in digital service design. These results suggest that the ability of chatbots to provide both emotionally rewarding and functionally efficient interactions is essential for enhancing satisfaction. In line with the literature on conversational AI (Lee, 2024; Mehra, 2021), the findings highlight that chatbots are no longer merely transactional tools but increasingly incorporate elements of empathy, personality, and social intelligence to elevate the quality of user interactions. In the mobile food delivery context, where service failures such as delayed or incorrect orders are common, such capabilities are particularly salient. Users' expectations extend beyond problem resolution to include speed, accuracy, transparency, and even enjoyment during interactions (Chandramouli, 2019; Lefebvre et al., 2024). This study demonstrates that chatbots capable of addressing both hedonic and pragmatic user needs are more likely to generate satisfaction, thereby fostering continuance intention.

Theoretical implications

This study makes several theoretical contributions by integrating and extending service recovery theory (Smith et al., 1999; Tax & Brown, 1998) and expectation-confirmation theory (Bhattacherjee, 2001; Oliver, 1980) in the context of AI-mediated complaint handling. Service recovery theory explains how users cognitively and affectively evaluate the chatbot-mediated recovery encounter following a service failure. In this context, hedonic experience reflects emotional and affective responses, while pragmatic experience captures functional and instrumental evaluations of the recovery effort itself. Expectation-confirmation theory, in turn, explains how these experiential evaluations translate into satisfaction and subsequent continuance intention. Satisfaction emerges as an evaluative judgment formed after users compare their expectations with the actual chatbot recovery experience and then functions as the primary driver of continuance intention.

Neither theory alone is sufficient to explain chatbot-based complaint handling. Service recovery theory focuses on post-failure evaluations but does not explicitly explain how these evaluations translate into long-term behavioral intentions. Conversely, expectation-confirmation theory explains satisfaction and continuance intention but remains largely silent on the experiential mechanisms through which service recovery interactions are perceived. By integrating these two perspectives, this study captures both the experiential formation of recovery perceptions through hedonic and pragmatic user experiences and the evaluative process through which these perceptions shape satisfaction and continuance intention.

The findings extend service recovery theory by demonstrating that recovery evaluations in AI-mediated contexts are not limited to outcome-based judgments but are strongly shaped by experiential dimensions of human–AI interaction. This enriches the theory with a process-oriented and interaction-focused perspective particularly relevant to AI-driven service encounters. The study also extends expectation-confirmation theory by empirically showing that satisfaction fully mediates the relationship between user experience and continuance intention in chatbot-based complaint handling. This positions satisfaction as a necessary evaluative mechanism rather than a parallel outcome, refining ECT by highlighting satisfaction as an essential bridge between experience and behavioral loyalty in AI-mediated recovery contexts. These findings suggest that in AI-driven service recovery, experiential factors precede evaluative judgments, which in turn drive behavioral intentions, thereby enriching both theories with an experiential and process-oriented perspective.

Managerial implications

From a managerial perspective, the findings highlight that improving chatbot functionality alone is insufficient to foster long-term user retention. Service providers should design chatbot-based complaint handling systems that simultaneously address pragmatic needs such as fast resolution, accuracy, and transparency and hedonic needs such as empathy, tone, and conversational quality. The dual-pathway influence of user experience on satisfaction underscores the importance of investing in both technical performance and conversational design.

The full mediating role of satisfaction suggests that positive experiences must be explicitly translated into satisfaction to influence continuance intention. Managers should therefore monitor not only interaction performance metrics but also user satisfaction following complaint resolution. Incorporating feedback mechanisms, post-interaction surveys, real-time satisfaction ratings, and sentiment analysis into chatbot systems can provide actionable insights into whether chatbot experiences are successfully converting into satisfaction and, ultimately, continued engagement.

Given that users' expectations extend beyond problem resolution to include emotional engagement and conversational quality, service providers should prioritize training AI models to recognize and respond to emotional cues, adopt empathetic language, and maintain a conversational tone that aligns with brand identity. In MFOAs, where service failures are common and competition is intense, such capabilities can serve as key differentiators in retaining users and building long-term loyalty.

Limitations and directions for future research agenda

Although this study provides valuable insights into how chatbot user experience influences satisfaction and continuance intention in MFOAs, several limitations should be acknowledged. This study did not control for service failure severity. Different types of complaints (e.g., food quality issues vs. delivery delays) may vary in severity, which could influence user satisfaction and continuance intention differently. While we measured general application usage recency, we did not specifically measure when participants filed their complaints through chatbots. Variations in complaint recency may introduce recall bias. Future research should explicitly measure both failure severity and complaint timing to control for these potential confounds.

The extremely high R^2 values observed for hedonic experience (0.916) and pragmatic experience (0.999) should be interpreted with caution, as they may partly reflect conceptual proximity inherent in hierarchical UX modeling and the use of repeated indicators in second-order formative constructs. Although discriminant validity and VIF results suggest no serious multicollinearity issues, future studies could replicate the model using alternative measurement specifications or higher-order modeling approaches.

We acknowledge that chatbot systems across platforms may rely on different underlying AI or LLM technologies, which were not directly observed or controlled in this study. As the focus of the present research is on users' perceived experience and post-usage evaluations rather than the technical characteristics of the chatbot systems, these potential differences were treated as unobserved contextual variance. However, such heterogeneity may introduce confounding effects that should be interpreted with caution. Future studies could address this limitation by focusing on a single platform, incorporating platform dummy variables, or explicitly measuring perceived chatbot intelligence or human-likeness.

Regarding measurement, satisfaction was assessed using items adapted from prior online shopping and IS continuance studies. Although these items have been widely validated, their adaptation to chatbot-mediated complaint handling may not fully capture context-specific aspects of post-failure satisfaction. Future research could develop and validate satisfaction measures tailored specifically to AI-based service recovery contexts.

The cross-sectional survey design limits the ability to draw definitive conclusions about causality. While our analysis supports the proposed relationships, longitudinal studies or experimental designs could offer a clearer picture of how user experience evolves over time and how it impacts satisfaction and continued use, especially in scenarios involving repeated service failures.

Sample composition may also affect the generalizability of the findings. Most participants were young adults, and women made up a larger proportion of the sample. Future research should include more diverse demographic groups to explore how factors such as age, occupation, and income influence the observed relationships. Investigating users from different cultural backgrounds could further clarify how expectations, trust, and engagement with chatbots vary across regions.

The focus on mobile food ordering limits the direct application of findings to other service contexts. Industries such as e-commerce, banking, healthcare, or telecommunications may have different patterns of service failures, complaint complexity, and user expectations. Extending the model to these sectors could help validate its broader applicability and refine theoretical frameworks related to expectation-confirmation and service recovery in AI-mediated interactions.

Relying exclusively on self-reported data introduces potential biases, including social desirability and recall errors. Future studies could address this limitation by combining survey responses with behavioral data, such as chatbot interaction logs, complaint resolution times, or actual usage patterns, to provide more objective assessments. Qualitative approaches, including interviews or experimental simulations, could also deepen understanding of the mechanisms through which pragmatic and hedonic experiences shape satisfaction and continuance intention.

It should also be acknowledged that Harman's single-factor test, while employed as a supplementary diagnostic, is increasingly regarded as insufficient for conclusively ruling out common method bias (Chin et al., 2012; Podsakoff et al., 2012). The test lacks statistical power and does not account for the multidimensional structure of constructs, limiting its diagnostic utility as a standalone remedy. Future studies should consider incorporating procedural remedies such as temporal separation of measurements, marker variable techniques, or the unmeasured latent method construct approach within CB-SEM frameworks where model identification permits.

Addressing these limitations in future research can improve both the theoretical understanding and practical design of chatbot-mediated service recovery, offering actionable insights for service providers seeking to enhance user satisfaction and long-term engagement.

Conclusion

This study demonstrates that user experience with chatbots in MFOAs plays an important foundational role in shaping user satisfaction, which in turn serves as a key driver of continuance intention. Hedonic and pragmatic aspects of user experience both contribute meaningfully to satisfaction, underscoring the value of designing chatbots that combine functional efficiency with engaging interaction qualities. However, the modest explanatory power observed for satisfaction indicates that user experience represents one rather than the sole determinant of satisfaction in chatbot-based complaint handling. Consistent with this, the direct effect of user experience on continuance intention was non-significant, highlighting the mediating role of satisfaction in translating experiential evaluations into sustained usage intentions.

By integrating expectation-confirmation and service recovery perspectives into an AI-mediated complaint handling context, this research extends these theoretical frameworks to digital service interactions. The findings suggest that effective chatbot design can support positive user evaluations by facilitating efficient problem resolution and fostering a sense of procedural adequacy, particularly in high-failure service environments such as food delivery. At the same time, the results indicate that satisfaction in AI-based service recovery is inherently multifaceted and likely influenced by additional outcome- and context-related factors beyond experiential perceptions alone. Accordingly, this study positions user experience as a critical but partial contributor to satisfaction, while demonstrating its central role in shaping continuance intention in chatbot-enabled service encounters.

Acknowledgements Taşkın Dirsehan is grateful to Erasmus University Rotterdam for covering the open-access article processing charge.

Author contribution Saule Yolcu: writing—review and editing, writing—original draft, validation, conceptualization, investigation. Alperen Şahin: writing—review and editing, writing—original draft, visualization, validation, project administration, methodology, investigation, data curation, conceptualization. Aslı Erdoğdu: writing—review and editing, writing—original draft, validation, data curation, investigation. Taşkın Dirsehan: supervision, writing—review and editing, writing—original draft, funding acquisition, validation, methodology, investigation, data curation, conceptualization.

Funding Erasmus Universiteit Rotterdam

Data availability Data will be made available on request.

Declarations

Ethical approval This study was reviewed and approved by the Ethics Committee of Istanbul Kent University, Social and Human Sciences Research Ethics Board (Approval Number: 2025/04, Date: 25.04.2025).

Competing interests The authors declare no competing interests.

References

Alabbas, A., & Alomar, K. (2025). A weighted composite metric for evaluating user experience in educational chatbots: Balancing usability, engagement, and effectiveness. *Future Internet, 17*(2), Article 64. https://doi.org/10.3390/fi17020064

Alotaibi, M., & Hidayat-ur-Rehman, I. (2025). An empirical analysis of user intention to use chatbots for airline tickets consultation. *Journal of Science and Technology Policy Management, 16*(1), 204–228.

Arif, M. A. I., Kabir, S., Islam, Z., Khan, M. F. H., Islam, I., & Islam, M. N. (2024). Usability evaluation of online food delivery applications in bangladesh: a comparative study. IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), Chennai, India, 2024. pp. 337-342. https://doi.org/10.1109/WIECON-ECE64149.2024.10914965

Aslam, F. (2023). The impact of artificial intelligence on chatbot technology: A study on the current advancements and leading innovations. *European Journal of Technology, 7*(3), 62–72.

Becker, J.-M., Klein, K., & Wetzels, M. (2012). Hierarchical latent variable models in PLS-SEM: Guidelines for using reflective-formative type models. *Long Range Planning, 45*(5–6), 359–394.

Bhattacharya, C., & Sinha, M. (2022). The role of artificial intelligence in banking for leveraging customer experience. *Australasian Business, Accounting and Finance Journal*. https://doi.org/10.14453/aabfj.v16i5.07

Bhattacherjee, A. (2001). Understanding information systems continuance: an expectation-confirmation model. *MIS quarterly, 25*(3), 351–370. https://doi.org/10.2307/3250921

Bhattaru, S., Goli, M., Swetha, T., Soujanya, R., & Jain, A. (2024). Customer service 2.0: the influence of chatbots and AI solutions. *MATEC Web of Conferences, 392*, 01041. https://doi.org/10.1051/matecconf/202439201041

Chagas, B. A., Pagano, A. S., Prates, R. O., Praes, E. C., Ferreguetti, K., Vaz, H., Reis, Z. S. N., Ribeiro, L. B., Ribeiro, A. L. P., & Pedroso, T. M. (2023). Evaluating user experience with a chatbot designed as a public health response to the COVID-19 pandemic in Brazil: Mixed methods study. *JMIR Human Factors, 10*(1), e43135.

Chandramouli, D. (2019). Consumer expectations on chatbots of food delivery apps. *Thinkquest International journal of technology and management, 4*(1), ISSN No: 2320 4214.

Cheng, Y., & Jiang, H. (2020). How do AI-driven chatbots impact user experience? Examining gratifications, perceived privacy risk, satisfaction, loyalty, and continued use. *Journal of Broadcasting & Electronic Media, 64*(4), 592–614.

Chin, W. W., Thatcher, J. B., & Wright, R. T. (2012). Assessing common method bias: Problems with the ULMC technique1. *MIS Quarterly, 36*(3), 1003–1019. https://doi.org/10.2307/41703491

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.

Chumkaew, S. (2023). The development of chatbot provided registration information services for students in distance learning. *ABAC Journal, 43*(4), 97–112. https://doi.org/10.59865/abacj.2023.55

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Dawar, S., Panwar, S., Dhaka, S., & Kudal, P. (2022). Antecedents and role of trust in chatbot use intentions: An Indian perspective. *Marketing i Menedžment Innovacij, 13*(4), 198–206.

Diederich, S., Brendel, A. B., Morana, S., & Kolbe, L. (2022). On the design of and interaction with conversational agents: An organizing and assessing review of human-computer interaction research. *Journal of the Association for Information Systems, 23*(1), 96–138.

Dirsehan, T., & Cankat, E. (2021). Role of mobile food-ordering applications in developing restaurants' brand satisfaction and loyalty in the pandemic period. *Journal of Retailing and Consumer Services, 62*, 102608.

Entenberg, G. A., Dosovitsky, G., Aghakhani, S., Mostovoy, K., Carre, N., Marshall, Z., Benfica, D., Mizrahi, S., Testerman, A., & Rousseau, A. (2023). User experience with a parenting chatbot micro intervention. *Frontiers in Digital Health, 4*, 989022. https://doi.org/10.3389/fdgth.2022.989022

Følstad, A., & Taylor, C. (2021). Investigating the user experience of customer service chatbot interaction: A framework for qualitative analysis of chatbot dialogues. *Quality and User Experience, 6*(1), 6.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50.

Fu, X.-m, Zhang, J.-h, & Chan, F. T. (2018). Determinants of loyalty to public transit: A model integrating satisfaction-loyalty theory and expectation-confirmation theory. *Transportation Research Part A, Policy and Practice, 113*, 476–490.

Fürst, A., Trißler, L., Friedrich, R., & Wirtz, J. (2025). Service recovery by AI or human agents: Do failure and strategy context matter? *Journal of Service Management, 36*(3), 390–418. https://doi.org/10.1108/josm-04-2024-0190

Ghali, Z. (2025). From an emotional experience of mobile food shopping to continued purchase intention: Moderating role of e-user expertise. *British Food Journal, 127*(1), 34–53. https://doi.org/10.1108/bfj-04-2024-0365

Hair, J. F. (2014). *A primer on partial least squares structural equation modeling (PLS-SEM)*. sage.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice, 19*(2), 139–152.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2–24. https://doi.org/10.1108/ebr-11-2018-0203

Haugeland, I. K. F., Følstad, A., Taylor, C., & Bjørkli, C. A. (2022). Understanding the user experience of customer service chatbots: An experimental study of chatbot interaction design. *International Journal of Human-Computer Studies, 161*, 102788. https://doi.org/10.1016/j.ijhcs.2022.102788

Haupt, M., Rozumowski, A., Freidank, J., & Haas, A. (2023). Seeking empathy or suggesting a solution? Effects of chatbot messages on service failure recovery. *Electronic Markets, 33*, 56.

Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management & Data Systems, 116*(1), 2–20.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*, 115–135.

Hinderks, A., & Thomaschewski, J. (2017). Design and evaluation of a short version of the user experience questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence, 4*(6), 103-108. https://doi.org/10.9781/ijimai.2017.09.001

Hoelter, J. W. (1983). The analysis of covariance structures: Goodness-of-fit indices. *Sociological Methods & Research, 11*(3), 325–344. https://doi.org/10.1177/0049124183011003003

Hsiao, C.-H., Chang, J.-J., & Tang, K.-Y. (2016). Exploring the influential factors in continuance usage of mobile social apps: Satisfaction, habit, and customer value perspectives. *Telematics and Informatics, 33*(2), 342–355.

Kang, B., & Hong, M. (2025). Development and evaluation of a mental health chatbot using ChatGPT 4.0: Mixed methods user experience study with Korean users. *JMIR Medical Informatics, 13*, e63538.

Kara, B. C., Şahin, A., & Dirsehan, T. (2025). BibexPy: Harmonizing the bibliometric symphony of Scopus and Web of Science. *SoftwareX*. https://doi.org/10.1016/j.softx.2025.102098

Kim, J. S., Kim, M., & Baek, T. H. (2025). Enhancing user experience with a generative AI chatbot. *International Journal of Human-Computer Interaction, 41*(1), 651–663. https://doi.org/10.1080/10447318.2024.2311971

Kim, M. (2023). Connecting artificial intelligence to value creation in services: Mechanism and implications. *Service Business, 17*(4), 851–878.

Kim, M. J. (2024). The effect of chatbot service quality on customer satisfaction and continuous use intention. *Journal of Korea Artificial Intelligence Association, 2*(1), 15–24.

Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of e-Collaboration (ijec), 11*(4), 1–10.

Kristjansson, E. A., Desrochers, A., & Zumbo, B. (2003). Designer's corner-translating and adapting measurement instruments for cross-linguistic and cross-cultural research: a guide for practitioners. *Canadian Journal of Nursing Research Archive*, *35*(2),127–142. PMID: 12908201

Kvale, K., Freddi, E., Hodnebrog, S., Sell, O. A., & Følstad, A. (2020). Understanding the user experience of customer service chatbots: what can we learn from customer satisfaction surveys? Chatbot Research and Design: 4th International Workshop, CONVERSATIONS 2020, Virtual Event, November 23–24, 2020, Revised Selected Papers p. 205-218. https://doi.org/10.1007/978-3-030-68288-0_14. Published: 23 November 2020 Publication History

Lee, W.-J. (2024). Service failure by human, service recovery by AI chatbot: The impact of justice, AI efficacy on recovery effort. *Virtual Economics, 7*(4), 48–63. https://doi.org/10.34021/ve.2024.07.04(3)

Lefebvre, S., Orlowski, M., & Boman, L. (2024). It's all your fault! Restaurant vs. platform blame attribution in food delivery service failures. *British Food Journal, 126*(8), 3037–3050. https://doi.org/10.1108/bfj-12-2023-1103

Ltifi, M. (2023). Trust in the chatbot: A semi-human relationship. *Future Business Journal, 9*(1), 109.

Macedo, P., Madeira, R. N., Santos, P. A., Mota, P., Alves, B., & Pereira, C. M. (2024). A conversational agent for empowering people with Parkinson's disease in exercising through motivation and support. *Applied Sciences, 15*(1), 223.

Malhotra, N. K. (2020). *Marketing research: An applied orientation*. Pearson.

McGorry, S. Y. (2000). Measurement in a cross-cultural environment: Survey translation issues. *Qualitative Market Research: An International Journal, 3*(2), 74–81. https://doi.org/10.1108/13522750010322070

Mehra, B. (2021). Chatbot personality preferences in Global South urban English speakers. *Social Sciences & Humanities Open, 3*(1), 100131.

Mejia, Y. Z. O. A., Navarro, C. A. D., Casildo-Bedón, N. E., & Pezo, Y. E. R. (2024). Influence of a chatbot based on a conversational agent on the adaptability of first-year students of a Peruvian private university. *Frontiers in Education, 9*, 1459303. https://doi.org/10.3389/feduc.2024.1459303

Melián-González, S., Gutiérrez-Taño, D., & Bulchand-Gidumal, J. (2021). Predicting the intentions to use chatbots for travel and tourism. *Current Issues in Tourism, 24*(2), 192–210.

Meng, H., Lu, X., & Xu, J. (2025). The impact of chatbot response strategies and emojis usage on customers' purchase intention: The mediating roles of psychological distance and performance expectancy. *Behavioral Sciences, 15*(2), 117.

Merkouris, S. S., Loram, G., Abdelrazek, M., Rodda, S. N., Ibrahim, A., Bonti, A., & Dowling, N. A. (2024). Improving the user experience of a gambling support and education website using a chatbot. *Universal Access in the Information Society, 23*(1), 213–225. https://doi.org/10.1007/s10209-022-00932-5

Merli, R., Preziosi, M., & Acampora, A. (2018). How do scholars approach the circular economy? A systematic literature review. *Journal of Cleaner Production, 178*, 703–722. https://doi.org/10.1016/j.jclepro.2017.12.112

Mustafa, H. (2024). Exploring the role of AI in customer service communication a case study on chatbots in delivery apps in UAE. *Balkan Social Science Review, 24*(24), 445–469. https://doi.org/10.46763/bssr242424445m

Nitzl, C., Roldan, J. L., & Cepeda, G. (2016). Mediation analysis in partial least squares path modeling: Helping researchers discuss more sophisticated models. *Industrial Management & Data Systems, 116*(9), 1849–1864.

Nunnally, J. C. (1978). An overview of psychological measurement. *Clinical diagnosis of mental disorders: A handbook*, pp. 97–146, Springer. https://doi.org/10.1007/978-1-4684-2490-4_4

Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research, 17*(4), 460–469.

Ozuem, W., Ranfagni, S., Willis, M., Salvietti, G., & Howell, K. (2024). Exploring the relationship between chatbots, service failure recovery and customer loyalty: A frustration–aggression perspective. *Psychology & Marketing, 41*(10), 2253–2273.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., & Brennan, S. E. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*. https://doi.org/10.1136/bmj.n71

Park, J. E., Fan, A., & Wu, L. (2025). Chatbots in complaint handling: The moderating role of humor. *International Journal of Contemporary Hospitality Management, 37*(3), 805–824.

Pezenka, I., Aunimo, L., Janous, G., & Dobrowsky, D. (2024). Emotionality in task-oriented chatbots–the effect of emotion expression on chatbot perception. *Communication Studies, 75*(6), 825–843. https://doi.org/10.1080/10510974.2024.2363259

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879. https://doi.org/10.1037/0021-9010.88.5.879

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology, 63*(1), 539–569. https://doi.org/10.1146/annurev-psych-120710-100452

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*(3), 879–891.

Rahmanti, A. R., Yang, H.-C., Bintoro, B. S., Nursetyo, A. A., Muhtar, M. S., Syed-Abdul, S., & Li, Y.-C. (2022). SlimMe, a chatbot with artificial empathy for personal weight management: System design and finding. *Frontiers in Nutrition, 9*, 870775.

Rajabi, E., George, A. N., & Kumar, K. (2024). The role of knowledge graphs in chatbots. *The Electronic Library, 42*(3), 483–497.

Rese, A., & Witthohn, L. (2025). Recovering customer satisfaction after a chatbot service failure–the effect of gender. *Journal of Retailing and Consumer Services, 84*, 104257.

Ringle, C. M., Wende, S., & Becker, J.-M. (2024). SmartPLS 4. In *SmartPLS*. https://www.smartpls.com. Accessed 15 May 2025

Şahin, A., Kara, B. C., & Dirsehan, T. (2025). LitOrganizer: Automating the process of data extraction and organization for scientific literature reviews. *SoftwareX, 30*, 102198. https://doi.org/10.1016/j.softx.2025.102198

Saklani, S., & Kala, D. (2024). Perception of GEN Z customers towards chatbots as service agents: A qualitative study in the Indian context. *Journal of Telecommunications and the Digital Economy, 12*(1), 356–376.

Sands, S., Campbell, C., Ferraro, C., & Plangger, K. (2022). Buffering B2B service failure: The role of customer engagement. *Industrial Marketing Management, 103*, 47–60. https://doi.org/10.1016/j.indmarman.2022.03.007

Sarstedt, M., Hair, J. F., Jr., Cheah, J.-H., Becker, J.-M., & Ringle, C. M. (2019). How to specify, estimate, and validate higher-order constructs in PLS-SEM. *Australasian Marketing Journal, 27*(3), 197–211.

Sarstedt, M., Hopkins, L., & Kuppelwieser, V. G. (2014). Partial least squares structural equation modeling (PLS-SEM): An emerging tool in business research. *European Business Review, 26*(2), 106–121.

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2021). Partial least squares structural equation modeling. *Handbook of market research* (pp. 587–632). Springer.Cham. https://doi.org/10.1007/978-3-319-57413-4_15

Shahzad, M. F., Xu, S., An, X., & Javed, I. (2024). Assessing the impact of AI-chatbot service quality on user e-brand loyalty through chatbot user trust, experience and electronic word of mouth. *Journal of Retailing and Consumer Services, 79*, 103867.

Shareef, F. (2024). Enhancing conversational AI with LLMs for customer support automation. 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS), Erode, India, 2024. pp. 239-244. https://doi.org/10.1109/ICSSAS64001.2024.10760403

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.

Sohail, M., Wang, F., Archer, N., Wang, W., & Yuan, Y. (2024). Lower than expected but still willing to use: User acceptance toward current intelligent conversational agents. *Information & Management, 61*(8), 104033.

Song, M., Xing, X., Duan, Y., & Mou, J. (2023). I can feel AI failure: The impact of service failure type and failure assessment on customer recovery expectation. *Industrial Management & Data Systems, 123*(12), 2949–2975.

Statista. (2025). *Consumer opinions on use of conversational AI for customer service in 2024*. Retrieved 30.03.2025 from https://www.statista.com/statistics/1538260/consumer-opinions-on-conversational-ai/. Accessed 30 Mar 2025

Stefanidi, E., Foukarakis, M., Arampatzis, D., Korozi, M., Leonidis, A., & Antona, M. (2019). ParlAmI: A multimodal approach for programming intelligent environments. *Technologies, 7*(1), 11.

Tax, S. S., & Brown, S. W. (1998). Recovering and learning from service failure. *MIT Sloan Management Review, 40*, 75-88. https://sloanreview.mit.edu/article/recovering-and-learning-from-service-failure/

Van Vaerenbergh, Y., Varga, D., De Keyser, A., & Orsingher, C. (2019). The service recovery journey: Conceptualization, integration, and directions for future research. *Journal of Service Research, 22*(2), 103–119.

Voorhees, C. M., Brady, M. K., Calantone, R., & Ramirez, E. (2016). Discriminant validity testing in marketing: An analysis, causes for concern, and proposed remedies. *Journal of the Academy of Marketing Science, 44*, 119–134.

Wang, S., Yan, Q., & Wang, L. (2025). Task-oriented vs. social-oriented: Chatbot communication styles in electronic commerce service recovery. *Electronic Commerce Research, 25*(3), 1793–1825. https://doi.org/10.1007/s10660-023-09741-1

Wang, Y.-L., & Lo, C.-W. (2025). The effects of response time on older and young adults' interaction experience with chatbot. *BMC Psychology, 13*(1), 150.

Wut, T. M., Chan, E.-h, & Wong, H.-m. (2024). Does fun matter? Using chatbots for customer services. *Informatics*. https://doi.org/10.3390/informatics11040094

Xing, Y. X. (2024). Exploring the use of ChatGPT in learning and instructing statistics and data analytics. *Teaching Statistics, 46*(2), 95–104. https://doi.org/10.1111/test.12367

Yoon, J., & Yu, H. (2022). Impact of customer experience on attitude and utilization intention of a restaurant-menu curation chatbot service. *Journal of Hospitality and Tourism Technology, 13*(3), 527–541.

Zhang, R. W., Liang, X., & Wu, S.-H. (2024). When chatbots fail: Exploring user coping following a chatbots-induced service failure. *Information Technology & People, 37*(8), 175–195.

Zhao, X., Lynch, J. G., Jr., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research, 37*(2), 197–206.