

A Study on the Continued Use of AI-Enabled mHealth Services Among Older Adults from a User Experience Perspective: The Moderating Effect of AITS

Xiang Wang¹, Jia Du¹, Jianjun Hou¹, Zhizheng Zhang¹, Shengzan Yan^{1,*}

¹School of Art and Design, Nanjing Institute of Technology, Nanjing, 211167 China

Abstract

INTRODUCTION: Against the backdrop of an accelerating aging population and a high prevalence of chronic diseases among the elderly, the imbalance between supply and demand for traditional healthcare resources has become increasingly prominent. While AI-enabled mobile health (mHealth) emerges as a vital solution to the challenges of elderly health management, critical issues such as the low Continued Usage Intention (CUI) among elderly users and the limitations of outdated research frameworks remain inadequately addressed.

OBJECTIVES: This study investigates the mechanisms underlying older adults' continued use of AI-enabled mHealth services, constructing and validating an adapted age-friendly technology acceptance framework that integrates the Expectation-Confirmation Model (ECM-ISC) with Quality of Experience (QoE), while examining the moderating role of AI Technology Services (AITS).

METHODS: Grounded in the ECM-ISC, this research introduces QoE as a core construct. User experience was decomposed into five dimensions: Sensory Experience (SE), Interaction Experience (IE), Functional Experience (FE), Emotional Experience (EE), and Trust Experience (TE). AITS was incorporated as a moderating variable to build an "Experience Dimensions–QoE–User Satisfaction (US)–CUI" framework. A questionnaire survey was conducted in East China in 2025, yielding 536 valid responses. Data were analyzed using Structural Equation Modeling (SEM) combined with moderation effect tests.

RESULTS: Empirical results demonstrate that: (1) All five experience dimensions significantly and positively influence QoE. IE, FE, and TE had the highest point estimates ($\beta = 0.24$), although pairwise tests showed the differences among the five dimensions ($\beta = 0.18\text{--}0.24$) were not statistically significant; EE had the lowest mean score ($M = 4.71$, $SD = 0.98$), indicating a current supply gap despite significant latent demand. (2) QoE acts as the core driver of CUI, with a path coefficient to US of 0.79 and a direct effect on CUI of 0.70—substantially higher than the direct effect of US itself ($\beta = 0.20$). This suggests elderly users rely more on rational judgments of overall service reliability than on transient emotional satisfaction when deciding to continue usage. (3) AITS significantly and positively moderates the relationships between SE, IE, FE, EE, and QoE; however, its moderating effect on the TE–QoE relationship was not significant ($B = 0.041$, 95% CI -0.004 to 0.086 , $p = 0.081$), so this study provides no evidence that AI features directly alleviate trust-related concerns.

CONCLUSION: This study contributes theoretically by adapting the traditional IS continuance model for an aging context, proposing an "adapted age-friendly technology acceptance framework," and delineating the experiential empowerment boundaries of AI technology. Practically, it proposes a tiered, differentiated support design strategy prioritizing "functionality and trust, optimizing interaction, and compensating emotion." These findings provide guidance for developing age-friendly digital health services that foster long-term engagement among the elderly, although longitudinal validation of actual usage behavior remains needed.

Keywords: AI-Enabled mHealth Services, Older Adults, User Experience, Quality of Experience, Continued Use Intention

Received on 29 July 2026, accepted on 29 September 2026, published on 20 October 2026

Copyright © 2026 Xiang Wang *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [CC BY-NC-SA 4.0](#), which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/eetpht.11.14224

*Corresponding author. Email: shengzanyan@njit.edu.cn

1. Introduction

Since the onset of the 21st century, global population aging has intensified at an unprecedented pace. Recent statistics from China indicate that the population aged 60 and above has reached 321.22 million, accounting for 22.86% of the national total, while those aged 65 and above stand at 223.09 million (15.87%) [1]. This demographic shift has precipitated a higher prevalence of chronic conditions among the elderly, resulting in a surge in healthcare demand that far exceeds that of other age cohorts [2]. Currently, over 180 million older adults in China live with chronic illnesses, with multimorbidity emerging as a particularly salient issue [3]. However, under the conventional healthcare paradigm, China faces persistent challenges regarding the uneven distribution, low utilization efficiency, and misallocation of medical resources, leaving the escalating healthcare needs of the elderly largely unmet [4].

In this context, mobile health has been recognized by the World Health Organization as a pivotal strategy for health management. Empirical studies substantiate that mHealth serves as a critical complement to chronic disease management [5], offering the potential to reduce costs, enhance care quality, and improve accessibility [6]. Specifically, mHealth facilitates health monitoring and disease management for older adults [7] and is widely regarded as a viable solution to the aging care crisis [8]. Moreover, the convergence of AI and mobile internet technologies presents unprecedented opportunities to rebalance medical supply and demand [9]. The integration of AI—encompassing deep learning, big data analytics, semantic understanding, and medical chatbots—into platforms such as “Left Hand Doctor,” Medwhat, and Babylon, has significantly optimized diagnostic efficiency and user experience, thereby expanding the reach of high-quality medical resources.

Despite these technological advancements, significant barriers remain for elderly users:

- **Low Continued Usage Intention:** Despite having the highest healthcare needs, older adults are not the primary users of mHealth services. As of December 2024, China had 418 million Internet healthcare users, accounting for 37.7% of all netizens; yet netizens aged 50 and above made up only 34.1% of the netizen population, and only 47.4% of netizens aged 60 and over were able to use mobile apps in elderly mode [10]. In the same period, adults aged 60 and above constituted 46.8% of the 301 million non-Internet users [10]. Historical data further reveals a stark disparity: in Q1 2021, users aged 40 and above accounted for a mere 5.4% of the mHealth user base, compared to over 80% represented by those aged 25-40 [11]. This indicates a considerable gap between service availability and actual adoption among the elderly.
- **Outdated Research Frameworks:** Although scholarly interest in AI-enabled mHealth is growing, the empirical evidence regarding its effectiveness in geriatric care remains contentious. Current research lacks frameworks

that effectively adapt to the cognitive and physical decline associated with aging [12]. While AI introduces novel functionalities, it also imposes new cognitive loads and usability challenges. Although seniors may acknowledge the utility of AI for convenience and health management, their acceptance is often conditional—constrained by algorithmic opacity, trust deficits, and interface complexity [13]. Furthermore, the absence of unified standards for age-friendly user experience (UX) design exacerbates these issues [14].

These gaps give rise to critical research questions: Does AI technology genuinely enhance or inadvertently hinder the user experience for older adults? What are the specific experiential dimensions involved in their interaction with AI-driven mHealth services? How is CUI formed, and which factors critically determine it?

Presently, research on AI-supported UX and continuance behavior remains fragmented. There is an urgent need for a systematic theoretical framework to elucidate these mechanisms. Accordingly, this study aims to investigate the formation mechanism of CUI among older users of AI-supported mHealth services, with a specific focus on deconstructing the key experiential factors that influence sustained engagement from a service experience perspective.

2. Literature Review

2.1 User Experience of AI-enabled mHealth Applications

Compared with traditional mHealth applications, AI-enabled platforms fundamentally transform the user experience through advanced interactive paradigms (Table 1) [15,16]. The primary modalities of interaction within these systems include medical chatbots, virtual assistants, and AI-enabled diagnostics [17,18]. Among these, chatbots represent the most prevalent interface for service delivery. Defined as “software programs that interact with users via natural language” [19,20], these agents utilize natural language processing (NLP) to comprehend user inputs and generate real-time, contextually appropriate responses through text or voice [21,22]. Recent AI-assisted inclusive interaction research shows that adaptive interfaces can improve accessibility while preserving usability across diverse cognitive profiles [23].

While AI-enabled healthcare chatbots possess the capacity to process vast linguistic datasets, significant limitations persist in their ability to fully comprehend nuanced human communication, particularly among aging populations [24,25]. These limitations manifest in four critical dimensions:

First, technological unfamiliarity. Compared to younger generations, older adults often exhibit lower digital literacy regarding AI-specific functions. This knowledge gap increases the cognitive barrier to entry, making the initial adoption of AI-enabled healthcare services more arduous.

Second, constrained interaction modalities. Current AI healthcare services predominantly rely on textual or vocal interfaces. However, age-related declines in visual acuity, auditory perception, or articulatory precision can severely impede interaction. For instance, presbycusis (age-related hearing loss) or dysarthria may prevent the system from accurately capturing or interpreting user commands, leading to communication breakdowns.

Third, limitations in information comprehension and processing. AI algorithms still struggle with the complexities of medical jargon, regional dialects, and diverse accents. Given that older users may describe symptoms using colloquialisms or non-standard terminology, the system's inability to accurately decode these inputs results in irrelevant or erroneous feedback, thereby frustrating the user.

Fourth, the issues of AI interpretability and trustworthiness. The vast majority of mainstream medical AI models—particularly those based on deep neural networks—

fundamentally retain a pronounced "black-box" nature. The inferential pathways between their inputs and outputs remain highly opaque, to the extent that even the developers themselves struggle to precisely trace the internal mechanisms by which the models arrive at specific conclusions. When applied to the elderly population, this technical characteristic gives rise to a particularly acute dilemma. Recent empirical research by Li et al. (2025) [26] provides direct experiential evidence for the trust-enhancing effect of Explainable AI (XAI) in elderly health information systems.

Collectively, these experiential disruptions and deficiencies increase the operational difficulty for elderly users and objectively degrade the QoE. Such impediments are likely to erode trust and diminish the perceived utility of the service, ultimately undermining older adults' intention to sustain usage of AI-enabled mHealth technologies.

Table 1. Comparison of User Experience Between AI-enabled mHealth Apps and Traditional mHealth Apps

Category	mHealth Applications Powered by AI Technology	Traditional mHealth Apps
Experience Type	Conversational AI Bot Experience	Search, Inquiry, and Other User Experiences
Functional Experience	AI Healthcare Services: Understanding user language and intent to quickly generate responses	Healthcare Services: Quick search, sorting, and display of existing data
Input Method	Multi-round text conversations or voice-assisted conversations, etc.	Keywords, etc.
Experience Benefits	Based on specific use cases, we build and train intelligent conversational bots, configure them with smart Q&A capabilities, and design multi-round conversations. This enables bots to assist human service providers, thereby improving user efficiency and satisfaction.	Providing Diverse Search Results
Presentation Services	Accurate delivery of text, links, tables, programs, and more	Fuzzy search for text, links, etc.

2.2 Current State of Research on mHealth

Contemporary research on mHealth has evolved along multiple trajectories, primarily converging on four key domains: (1) Industry Development and Policy, (2) Technical Design and Implementation, (3) Service Intervention Effectiveness, and (4) Adoption and Continuance Behavior.

First, from the perspective of industrial development, scholarly attention has focused on market dynamics, promotional strategies, and regulatory frameworks. Studies emphasize that stakeholders must adhere to FDA guidelines to standardize app development, particularly suggesting that online platforms foster deeper collaborations with established

medical device manufacturers to navigate environments characterized by radical innovation [27-28]. Furthermore, Jogova et al. (2019) [29] highlight the regulatory lag in regions like Canada, where the lack of specific mHealth legislation and a persistent conceptualization of medical devices as purely hardware entities may stifle innovation.

Second, from a technical and design perspective, research emphasizes the architecture and user interface (UI) optimization of mHealth platforms. Zhang & Liu (2022) [30] conducted technical analyses of mobile devices, implementing user-experience-oriented designs to enhance usability. Extending this, Zhang et al. (2024) [31] addressed the issue of power asymmetry between physicians and

patients in online consultations, proposing UI design solutions that mitigate this imbalance and contribute novel insights to telemedicine design.

Third, from the perspective of service intervention, scholars have evaluated the efficacy of mHealth in clinical settings. Kitsiou et al. (2025) [32], through a randomized controlled trial, demonstrated that patient-centered mHealth interventions are feasible, acceptable, and safe, and can effectively improve self-care among patients with chronic heart failure, while also enhancing their health beliefs and self-efficacy. Similarly, Ruiz-Leon et al. (2025) [33] showed that mHealth interventions for diabetic patients stimulate intrinsic motivation, promote healthier lifestyles, and strengthen daily self-management skills. Furthermore, Candelaria et al. (2025) [34] found that personalized recommendation services enhance self-care among elderly individuals at high risk of cardiovascular disease.

Fourth, concerning adoption and continuance behavior, research explores the technological, environmental, and cognitive determinants of usage. Wang et al. (2023) [35] confirmed that attitudes toward use and perceived ease of use positively influence older adults' behavioral intentions, mediated by perceived usefulness. Building on the UTAUT model, Wang et al. (2023) [36] identified that performance expectancy, effort expectancy, and social influence positively correlate with usage intentions, whereas perceived cost and perceived risk act as significant barriers.

Despite these significant advances, the literature on CUI in mHealth remains fragmented and faces two principal limitations:

First, there is a paucity of research examining CUI through the lens of UX. While existing studies have identified various technical and social factors, they often overlook the central role of subjective experiential perception. Users are the ultimate agents of mHealth adoption; however, the underlying mechanisms by which the QoE—particularly the emotional and cognitive responses during interaction—shapes long-term CUI among older adults remain insufficiently explored.

Second, the transformative impact of AI on the user experience has been largely neglected. With the deep integration of AI into mHealth, the technology no longer merely serves as a tool but fundamentally alters the interaction paradigm. Current research lacks an in-depth analysis of how AI acts as a core driver or moderator in reshaping the UX process and quality. Specifically, there is an urgent need to understand how AI-specific attributes (e.g., algorithmic transparency, conversational fluidity, and personalization) influence the formation of CUI among the elderly.

Therefore, to bridge these gaps, this study seeks to establish a comprehensive framework that integrates AI-driven UX factors into the formation mechanism of CUI for older adults.

2.3 Expectation-Confirmation Model of Information Systems Continuance

Research on CUI originates from the scholarly endeavor to understand users' post-adoption behavior within the realm of Information Systems (IS). Grounded in the Expectation Confirmation Theory, Bhattacherjee (2001) [37] pioneered the ECM-ISC, as illustrated in Figure 1. This parsimonious yet robust model defines CUI as the dependent variable reflecting a user's subjective propensity to persist in using an information system.

The ECM-ISC posits that continued usage intention is fundamentally determined by two key post-usage cognitive constructs: US and Perceived Usefulness (PU). Specifically, the model proposes that users' decisions to continue usage are driven by their satisfaction with prior experiences, which is in turn shaped by the confirmation of initial expectations and the perceived utility of the service. Due to its strong explanatory power regarding the psychological mechanisms underlying sustained engagement, the ECM-ISC has become a cornerstone theoretical framework for investigating user retention in the IS discipline, including in the context of mHealth services.

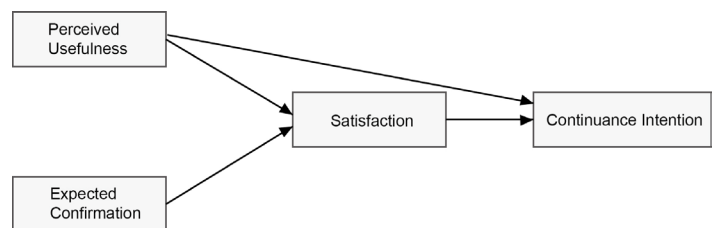


Figure 1. ECM-ISC

2.4 Concepts Related to User Experience Factors in AI-enabled mHealth Applications

2.4.1 Quality of Experience

The concept of QoE originated in telecommunications engineering, serving as a holistic metric to evaluate the end-to-end perceived performance of services from the user's perspective. Initially, QoE was largely conceptualized as an extension of Quality of Service (QoS), with scholarly focus centered on how network transmission parameters influenced user perception [38]. However, with the deepening of interdisciplinary research, the scope of QoE has transcended purely technical metrics to prioritize the measurement of subjective acceptability and overall perceived value. The International Telecommunication Union (ITU) provides the authoritative definition of QoE as “the overall acceptability of an application or service, as perceived subjectively by the end-user.”

Within the domain of AI-enabled mHealth, QoE emerges as a pivotal construct for assessing the extent to which digital health services fulfill the needs of elderly patients. Moving

beyond rudimentary usability metrics, QoE captures the multifaceted nature of human-AI interaction.

Drawing upon the systemic framework proposed by Akter et al. (2020) [39], this study adopts its multidimensional perspective while specifying **QoE as a holistic evaluative judgment formed by elderly users after interacting with an AI-enabled mHealth service. **Rather than operationalizing QoE as a higher-order factor whose indicators are the five dimensions, the present model treats QoE as a distinct reflective latent construct measured by three global evaluative items (QoE1–QoE3), and models the five experiential dimensions—SE, IE, FE, EE, and TE—as its key antecedents rather than as its measurement indicators (H1–H5). This conceptualization not only reflects the operational efficacy of the system but also embodies the comprehensive fulfillment of elderly users' physiological, psychological, and socio-emotional needs within an intelligent healthcare ecosystem.

2.4.2 Sensory Experience

SE refers to the immediate perceptual responses users derive from interacting with a system through multiple sensory channels—primarily vision, audition, and touch. Within the field of Human-Computer Interaction (HCI), the core value of SE lies in its capacity to minimize cognitive load and enhance the efficiency and accuracy of information transmission, achieved through sensory design aligned with users' cognitive schemas and physiological capabilities [40]. With the integration of artificial intelligence, wearable sensors, and advanced haptic feedback technologies, mHealth applications are evolving from purely function-oriented tools into intelligent ecosystems that support rich, multi-channel interactions.

In the context of this study, SE specifically denotes the holistic perceptual quality perceived by elderly users when interfacing with AI-enabled mHealth applications. Given the age-related decline in sensory acuity (e.g., presbyopia, presbycusis), the design of SE becomes critically important. It encompasses visual interface design (e.g., font legibility, icon clarity, and color contrast), auditory feedback (e.g., clarity and speed of voice announcements), haptic responses (e.g., vibration intensity and tactile cues), and the overall coherence of multimodal information presentation.

From an engineering perspective, SE is operationalized through measurable technical parameters such as interface rendering latency, color contrast ratios, font scalability, audio bitrate for voice feedback, and haptic vibration duration. These parameters directly determine whether the system aligns with age-related sensory declines, forming a clear "technical feature → sensory perception → user evaluation" chain. These thresholds are grounded in accessibility and human-centred design standards [79–82].

2.4.3 Interactive Experience

IE refers to the holistic perception and cognitive evaluation formed by users during two-way, dynamic interactions with a system. Within the field of human-computer interaction, its core objective is to reduce operational resistance through

interaction design aligned with users' mental models, thereby enhancing task completion efficiency and subjective satisfaction [41]. With the advancement of adaptive user interfaces and context-aware technologies, mHealth applications are evolving from one-way information display tools into intelligent agents capable of conducting natural, continuous dialogues with users.

In the context of this study, IE specifically denotes the subjective evaluation by older users regarding system response efficiency, operational fault tolerance, navigation clarity, and the naturalness of interaction when performing tasks via AI-enabled mHealth applications. Specifically, it encompasses the fluency of guidance based on multi-turn dialogues, as well as context-aware, non-intrusive interaction design aimed at enhancing older users' sense of operational efficacy and control.

Technically, IE is driven by speech recognition accuracy (particularly for regional dialects and age-related dysarthria), natural language understanding (NLU) confidence scores, multi-turn dialogue state tracking, and error-recovery mechanisms such as undo/redo stacks. These engineering characteristics directly shape the "interaction fluency → perceived ease of use → continued usage intention" pathway. Reference benchmarks for these metrics are given in [79,81].

2.4.4 Functional Experience

FE refers to the overall perception and utility assessment formed by users regarding the practicality, effectiveness, completeness, and reliability of a system's core functions during actual use. Within technology acceptance models and information systems success theories, FE is widely recognized as a decisive foundation influencing users' adoption willingness and continued usage behavior [42,43]. With the deepening application of artificial intelligence algorithms in the healthcare sector, the functional value of mHealth is evolving from simple data recording toward complex decision-making support.

In the context of this study, FE specifically denotes the comprehensive perception of elderly users regarding the application's capability to address actual health issues. This encompasses the accuracy of auxiliary diagnosis, the continuity of health management, the adaptability of services, and the precision of resource matching.

From an engineering standpoint, FE is determined by recommendation algorithm accuracy, diagnostic model sensitivity/specificity, personalization engines based on user health profiles, and system uptime/reliability (SLA metrics). These technical metrics form the foundation of the "functional reliability → perceived usefulness → QoE" chain. Reference benchmarks are provided in [82,83].

2.4.5 Emotional Experience

EE pertains to the emotional responses, psychological resonance, and subjective well-being derived by users during their interactions with a system [44]. Serving as the critical link between cognition and behavior within the user experience framework, positive EE has been empirically

demonstrated to significantly enhance users' sense of identification with the system and their usage retention [45].

In the context of this study, EE specifically denotes the positive psychological evaluations generated among older users when interacting with AI-enabled mHealth applications, stemming from the system's empathetic design, anxiety-alleviation mechanisms, and positive reinforcement strategies. This encompasses the system's progressive disclosure strategies when handling sensitive medical information, stress-reduction designs during waiting or rehabilitation phases (such as VR-guided interventions), and affirmative feedback delivered through achievement mechanisms.

Collectively, these elements aim to mitigate the perceived "coldness" of technology and foster a perceived emotional connection with older users. It should be noted that this study's authors interpret such connection as technologically simulated rather than a replacement for genuine human empathy; this interpretation extends beyond the direct questionnaire evidence and should be validated in future qualitative research. On the engineering side, the affective counterpart of EE would be assessed through metrics such as automatic emotion-recognition accuracy, conversational empathy/agreement scores, and the alignment of response sentiment with the user's affective state; as with the engineering parameters in the other subsections, these metrics were not objectively measured in this study. The engineering field of affective computing and its associated metrics is introduced in [84].

2.4.6 Trust Experience

TE refers to users' comprehensive psychological perception of a system's safety, reliability, and ethical compliance within the high-stakes domain of healthcare. It encompasses multiple dimensions, including data privacy protection, algorithmic transparency, and human oversight mechanisms. Given the irreversible nature of medical decisions, users exhibit minimal tolerance for AI-induced errors; consequently, trust-building constitutes the cornerstone of sustainable development for mHealth applications [46].

In the context of this study, TE specifically denotes elderly users' overall confidence in AI-enabled mHealth applications regarding privacy protection, algorithmic transparency, and human-machine collaboration mechanisms. In light of the irreversibility of medical decisions, trust-building relies on a dual safeguard mechanism of "technological transparency plus human fallback." Specifically, the system must explicitly disclose the rationale behind its algorithmic outputs (i.e., explainable AI) and provide a clearly defined pathway for human intervention during critical junctures. This dual mechanism aims to alleviate users' apprehensions regarding "black-box" algorithms and mitigate concerns about algorithmic opacity.

Architecturally, TE is built upon explainable AI (XAI) modules that generate plain-language rationales for outputs, differential privacy mechanisms for data protection, audit logging for algorithmic decisions, and human-in-the-loop fallback interfaces. These engineering components constitute the "transparency & safety → trust → continuance" pathway,

which is critical for elderly users in high-stakes healthcare contexts. These components are discussed within the interpretability and AI-risk standards literature [83,86].

For analytical clarity, the engineering parameters listed in Sections 2.4.2–2.4.6 (e.g., rendering latency, contrast ratios, speech-recognition accuracy, diagnostic sensitivity/specificity, explainable-AI modules, and audit logging) are conceptual mappings between technical design levers and the corresponding experience dimensions, drawn from the accessibility and AI-risk standards cited above. They were not objectively measured in the present study: all 32 empirical items in Table 2 capture users' subjective perceptions. The "technical feature → user evaluation → engineering implication" chains are therefore proposed theoretical links rather than relationships directly tested here; measurable engineering targets derived from these mappings are revisited in Table 15 as propositions for future implementation and validation.

3. Research Methods

3.1 Research Hypotheses

3.1.1 The Impact of SE on QoE

Extensive literature confirms that visual aesthetics constitute a core determinant of user experience, capable of eliciting positive affective states such as satisfaction, enjoyment, and pleasure [47]. High-quality SE not only enhances the initial appeal of a system to new users but also serves as a critical mechanism for user retention and the optimization of overall QoE. Within the specific context of mHealth applications targeting elderly users, SE is manifested through interface fluidity, chromatic harmony, and layout logic. These visual-organizational elements directly shape older users' intuitive perceptions and act as antecedent variables influencing their QoE. Accordingly, we propose H1:

H1: Older users' perceived SE of mHealth applications has a significant positive impact on their QoE.

3.1.2 The Impact of IE on QoE

Zhao and Xie (2019) [48] assert that high-quality interpersonal interactions effectively satisfy users' perceived social benefits and their needs for respect and care, thereby enhancing satisfaction. Further empirical evidence from Poushneh & Vasquez (2017) [49] indicates that Augmented Reality (AR) interaction experiences significantly shape user experience by influencing multidimensional perceptions of product quality, which in turn affects satisfaction and purchase intention. In the context of this study, IE refers specifically to the human-computer interaction quality perceived by older users when operating mHealth applications. Drawing on this empirical support, it is posited that if applications are perceived as easy to learn, engaging, and highly error-tolerant at the interaction level, this will enhance older users' overall QoE. Thus, H2 is proposed: Older users' IE of mHealth applications has a significant positive impact on their QoE.

3.1.3 The Impact of FE on QoE

Existing empirical research underscores that FE is a core driver of platform attachment and CUI [50]. As a quintessential digital health tool, the core value of mobile health applications lies in their capacity to effectively address fundamental elderly needs in health management, medical consultation, and medication adherence [51]. For older users, the practicality, usability, and fault tolerance of these features directly determine the smoothness of achieving health management goals. Based on this premise, this study proposes H3: Older users' FE toward mHealth applications has a significant positive impact on their QoE.

3.1.4 The Impact of EE on QoE

Research by Domínguez-Quintero et al. (2021) [52] confirms that experiential authenticity positively influences QoE. Within the mobile health context, if older users experience positive emotions—such as care, support, and security—this can alleviate anxieties stemming from health issues and technology use, thereby fostering trust and reliance on the service. Consequently, this study infers that the emotional satisfaction and affective feedback derived from using these applications will exert a positive enhancing effect on QoE. Hence, H4 is proposed: Older users' EE of mHealth applications has a significant positive impact on their QoE.

3.1.5 The Impact of TE on QoE

Shen et al. (2025) [53] found that the anthropomorphism of virtual digital avatars positively impacts tourists' QoE. In the healthcare domain, TE—defined as users' trust in the system—constitutes a psychological reliance formed through a comprehensive assessment of the system's reliability, security, and promise-keeping ability. Given that mHealth involves sensitive physiological data and life-critical decisions, trust forms the cornerstone of user adoption. For older users, a high level of TE enhances their sense of control and security, translating into superior experience ratings. Therefore, we propose H5: Older users' TE toward mHealth applications has a significant positive impact on QoE.

3.1.6 The Impact of QoE on US and CUI

Deng et al. (2010) [54] established that the UX of information technology positively influences both US and CUI. In this study, QoE represents the intrinsic subjective reaction and overall evaluative judgment formed by elderly users regarding the application system's attributes and service delivery. It is inferred that when older users perceive a high QoE, this positive assessment of exceptional service will significantly enhance their satisfaction and recognition of the health benefits provided. This positive US and value perception will, in turn, become key drivers of long-term usage intention. Thus, the following hypotheses are proposed:

H6: Older users' QoE will have a significant positive effect on US.

H7: Older users' QoE will have a significant positive effect on CUI.

3.1.7 The Impact of US on CUI

In information systems and consumer behavior research, US is a pivotal psychological variable reflecting users' holistic evaluation after a period of actual use. Sharabati et al. (2022) [55] confirmed that TikTok users' satisfaction positively impacts their intention to continue using the platform. In this context, US refers to the level of contentment formed by older users based on their practical experiences with mobile health applications. Furthermore, according to the Information Systems Continuance Model, CUI is significantly determined by US. Therefore, we propose H8: Older adults' US with mobile health apps will have a significant positive effect on their CUI.

3.1.8 The Moderating Role of AITS on User Experience

AI technology significantly enhances users' operational convenience and psychological affinity by improving service precision (e.g., personalized health recommendations) and interaction naturalness (e.g., optimized voice interaction) [56]. For older users, AI-enabled health decision support (e.g., disease risk prediction) reinforces their perceived value across multiple experiential dimensions [57], specifically sensory [58], interactive [59], functional [60], emotional [61], and trust-based [62]. This perceptual enhancement suggests that AITS does not operate in isolation; rather, it acts as a critical moderating variable that strengthens or weakens the pathways through which specific experiential dimensions influence overall QoE. Design-driven AI healthcare research similarly emphasizes that human-centered interaction and trust are essential when translating technical capability into better user experience [63].

Based on this rationale, this study proposes the following set of moderating effect hypotheses:

H9a: AITS moderates the relationship between SE and QoE.

H9b: AITS moderates the relationship between IE and QoE.

H9c: AITS moderates the relationship between FE and QoE.

H9d: AITS moderates the relationship between EE and QoE.

H9e: AITS moderates the relationship between TE and QoE.

3.2 Research Model

This study aims to investigate the factors influencing the sustained use of mobile health applications by older adults in the context of AI technology. Based on a review of the literature, theoretical reasoning, and the aforementioned hypotheses, a research framework is proposed, as shown in Figure 2.

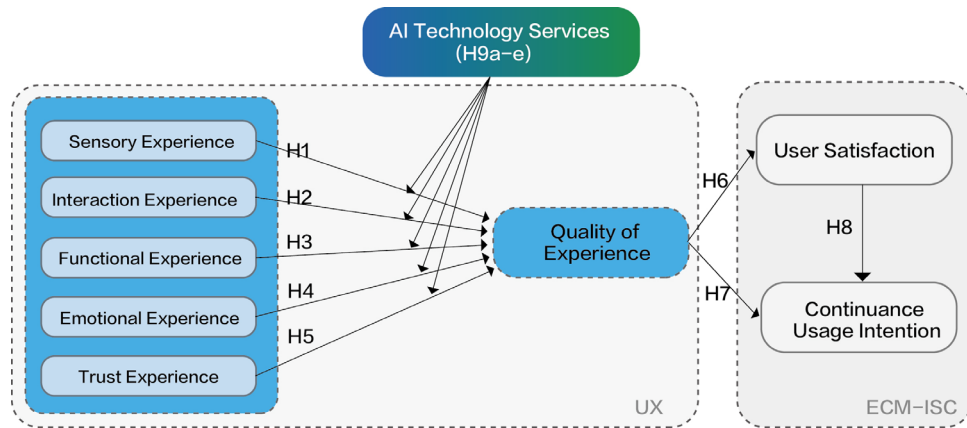


Figure 2. Research Model

3.3. Variable Definitions and Measurement

Based on the relevant literature, this study identified nine constructs: SE, IE, FE, EE, TE, QoE, US, CUI, and AITS. All items were measured using a 7-point Likert scale (1 =

strongly disagree, 7 = strongly agree). Definitions and sample items are shown in Table 2. AITS is operationalized as older users' perceived capacity of the AI service (a user attitude toward helpfulness, ease, and reliance), rather than as an objective characteristic of the service; objective AI service parameters would require separate engineering measurement.

Table 2. Definitions and Measurement of Research Variables

Construct	Definition of Constructs	Evaluation Items	Reference
SE	The immediate perceptual and holistic experience derived by elderly users through multiple sensory channels (sight, hearing, touch, etc.) when interacting with AI-enabled healthcare applications. This encompasses visual interface design, auditory feedback, tactile responses, and the presentation of multimodal information.	SE1.The visual design of this mHealth app's interface (e.g., font size, color contrast) is clear and comfortable. SE2.The app's auditory feedback (e.g., voice announcements, alert tones) is clear and easy to understand. SE3.The app's operational responses (e.g., vibration, haptic feedback) are timely and appropriate. SE4.The app's overall layout and information presentation align with my cognitive habits.	[64,65]
IE	The holistic perception formed by elderly users during two-way, dynamic interactions. It focuses on system responsiveness, operational fault tolerance, navigational clarity, interaction naturalness, and the fluency of multi-turn guidance, aiming to enhance users' sense of operational efficacy and control.	IE1.I find this app's controls simple, easy to learn, and user-friendly. IE2.This app allows me to easily undo or correct mistakes during operation. IE3.The app's multimodal interaction methods (e.g., voice, text) are natural and fluid. IE4.The app's navigation logic is clear, allowing me to easily locate desired features. IE5.The app's AI conversations and responses are accurate, reliable, and the interaction process is smooth.	[64,66]

FE	The capability of the app to address real-world health issues for older users, encompassing their overall perception and assessment regarding diagnostic accuracy, continuity of health management, service suitability, and the precision of resource matching.	FE1.The health information and services provided by this app are highly aligned with my actual needs. FE2.The app's AI-enabled diagnosis and health recommendations are accurate, reliable, and helpful. FE3.The app's health management services (e.g., chronic disease monitoring, medication reminders) are comprehensive and practical. FE4.This app effectively helps me identify suitable healthcare resources (e.g., doctors, hospitals).	[64,67]
EE	The positive psychological evaluations derived by older users, stemming from the system's empathetic design, anxiety-alleviation mechanisms, and positive reinforcement. This manifests as a sense of being cared for, a sense of security, enjoyment, and a mitigation of the perceived "coldness" of technology.	EE1.When using this app, I feel understood and cared for. EE2.Using this mobile health app does not induce anxiety or other negative emotions. EE3.While using this app, I feel relaxed and happy rather than anxious. EE4.I feel that the app's interactive design is warm and welcoming, rather than cold and mechanical.	[44,64]
TE	A comprehensive psychological assessment by older users regarding privacy protection, the scientific validity of diagnoses, and the limitations of AI within the context of AI-enabled mHealth. It reflects users' overall confidence in the system's trustworthiness.	TE1.I believe this app can effectively protect my personal privacy and health data security. TE2.I believe the app's AI diagnoses and recommendations are scientifically grounded, accurate, and authoritative. TE3.This app clearly informs me of the capabilities and limitations of its AI services.	[64,68]
QoE	The holistic subjective evaluation of overall service quality formed by older users regarding the objective attributes and service delivery of AI-enabled mHealth applications.	QoE1.Overall, I had a high-quality experience using this app. QoE2.The user experience of this app met my health management needs. QoE3.This app consistently delivers a high-quality experience across the health services I use.	[69,70]
US	The overall satisfaction of older users with the service features and user experience based on actual usage. It serves as a key mediating variable linking experience quality to continued usage intention.	US1.I am satisfied with the service features and experience provided by this app. US2.This app met my expectations for digital health services. US3.Overall, I am satisfied with my experience using this app.	[71,72]

CUI	The overall behavioral tendencies of older users, including intentions to continue usage, recommendation to others, and exploration of advanced features.	CUI1.I plan to continue using this mHealth app in the future.	
		CUI2.I am willing to try more AI-enabled health services offered by this app.	[73,74]
		CUI3.I would recommend this mHealth app to others.	
AITS	Older users' perceptions of the practicality and ease of use of AI technology, as well as their willingness to utilize it, reflecting the direct utility and value that AI functionality provides.	AITS1.The AI-enabled features in this app (e.g., smart consultations, health recommendations) are very helpful to me.	
		AITS2.AI technology makes this app easier to use and provides a more user-friendly experience.	[75,76]
		AITS3.I am willing to rely on the AI-enabled features in this app when managing my health.	

4. Empirical Analysis

4.1. Questionnaire Design

This study adopts a structured questionnaire survey approach. Following the recommendations of Oksenberg & Kalton (1991) [77], the required sample size for a pilot test typically ranges from 25 to 75 respondents. To this end, a preliminary field survey was conducted in Jiangsu Province from June 3 to 10, 2025, during which 100 paper-based questionnaires were administered. After excluding invalid responses characterized by logical inconsistencies or monotonic response patterns (e.g., excessive consecutive selection of the same option), 73 valid questionnaires were retained, satisfying and exceeding the aforementioned sample size criteria. The instrument was developed in Chinese. Items adapted from English-language sources were translated and back-translated by two bilingual researchers following the standard back-translation procedure [85], and the wording was then reviewed by an expert panel comprising two HCI/gerontechnology researchers, one geriatrician, and two senior mHealth practitioners. The panel-level content validity index (S-CVI/Ave) was 0.92 [87]. During the pilot, reliability analysis, item analysis, and unidimensionality testing were used to evaluate the scale; three items were revised—technical terms such as "multimodal" were replaced with plain language, one TE item was reworded to avoid a double-barreled statement, and AITS3 was adjusted to clearly separate reliance from frequency. The median completion time was about 11 minutes, confirming the feasibility and measurement quality of the instrument for the formal survey. Following the confirmation of questionnaire quality through the pilot test, the formal survey was conducted from July 16 to 28, 2025, targeting the East China region (primarily Jiangsu Province, Zhejiang Province, and Shanghai). Utilizing a convenience sampling strategy, the online questionnaire was distributed via the "Wenjuanxing"

(Questionnaire Star) platform. Eligibility required being aged 60 years or older and having used an AI-enabled mHealth application for at least three months; consequently, no non-user category was included, by design. Respondents were instructed to answer all experience items with reference to the AI-enabled mHealth app they used most frequently. Of the 536 valid responses, 79.1% were self-completed and 20.9% were completed with assistance from family members or community volunteers, who read the questions and response options aloud without interpreting or guiding the answers. A total of 600 questionnaires were initiated, and 64 were excluded for the following reasons: 24 were completed in less than two minutes (speeding), 18 showed straight-lining or monotonic response patterns, 13 failed the logical-consistency screening item, and 9 had more than 10% missing values. This yielded 536 valid responses (a valid response rate of 89.33%). All participants gave informed consent after being fully informed of the survey's purpose. The protocol was approved by the Institutional Review Board (IRB) of the School of Art and Design at Nanjing Institute of Technology (Approval No. NJIT-SADE-20250412), and strict data protection procedures were implemented, including anonymization and encrypted storage.

The final questionnaire comprised 38 items in total, including 32 core measurement items (distributed across 9 constructs) and 6 demographic/screening questions (gender, age, education level, monthly income, duration of use, and a screening item for logical consistency). The sample size of 536 valid responses meets the minimum requirement for Maximum Likelihood (ML) estimation proposed by Jackson (2003) [78], which stipulates a parameter-to-sample-size ratio (p:n) of at least 1:10. Statistical analyses were performed using SPSS 22.0 and AMOS 17.0.

4.2. Descriptive Statistics

Based on the statistical analysis of the valid questionnaires, the distribution of demographic characteristics of the sample is shown in Table 3. In terms of gender, males accounted for 52.20% and females for 47.80%; in terms of age, the 60–65 group had the highest proportion (47.20%), followed by the 66–70 group (27.80%) and those aged 71 and older (25.00%); regarding educational attainment, the majority had a high school education or below (69.20%), with associate's and bachelor's degrees accounting for 18.10% and master's degrees and above for 12.70%; regarding monthly income, the 3,001–5,000 yuan range had the largest share (35.30%), followed by less than 3,000 yuan (34.70%) and more than 5,000 yuan (30.00%); regarding duration of use, the largest group had used AI medical applications for 1–3 years (40.30%), while those with less than 1 year and more than 3

years accounted for 23.70% and 36.00%. The apps respondents anchored on were distributed as follows: Ping An Health 27.4%, JD Health 20.9%, Alibaba Health/YiLu 11.6%, WeDoctor 10.8%, Good Doctor Online 8.6%, tertiary or community hospital-affiliated apps 13.2%, and other specialist chronic-disease apps 7.5%. The most commonly used AI functions were AI conversational consultation (78.4%), personalized health recommendations (71.2%), medication and health reminders (68.7%), AI risk assessment and report interpretation (54.3%), and voice or dialect interaction (41.6%). Overall, the sample covers the principal demographic strata of older AI-enabled mHealth users in the surveyed region. Because it is a non-probability convenience sample, it should not be interpreted as a statistically representative sample, and generalization beyond the region should be made with caution. The means and standard deviations of the nine constructs are reported in Table 4.

Table 3. Descriptive Statistics for Demographic Variables

Variables	Options	Frequency	Percentage	Cumulative Percentage
Gender	Male	280	52.20	52.20
	Female	256	47.80	100.00
Age	60–65	253	47.20	47.20
	66–70	149	27.80	75.00
	71 and older	134	25.00	100.00
Education Level	High school or below	371	69.20	69.20
	College or undergraduate	97	18.10	87.30
	Master's degree or higher	68	12.70	100.00
Monthly Income	Less than 3,000 yuan	186	34.70	34.70
	3,001–5,000 yuan	189	35.30	70.00
	More than 5,000 yuan	161	30.00	100.00
Duration of Use	Less than 1 year	127	23.70	23.70
	1–3 years	216	40.30	64.00
	More than 3 years	193	36.00	100.00

Table 4. Means and Standard Deviations of Constructs

Construct	Mean	SD
SE	5.11	0.99
IE	4.98	0.95
FE	5.03	0.99
EE	4.71	0.98
TE	4.94	0.98
QoE	5.08	1.04
US	5.11	1.03
CUI	5.11	1.03
AITs	4.80	1.01

4.3. Reliability Analysis

In social science research, Cronbach's α coefficient is the most widely adopted metric for evaluating reliability. Conventionally, the interpretation of reliability coefficients follows established thresholds: a value above 0.9 indicates excellent reliability; 0.8–0.9 reflects good reliability; 0.7–0.8 suggests acceptable reliability; 0.6–0.7 is considered marginally acceptable; and values below 0.6 indicate insufficient reliability, necessitating item revision.

As presented in Table 5, the Cronbach's α coefficients for all constructs in this study exceed 0.8, demonstrating that the measurement instrument possesses good to excellent reliability. Furthermore, an examination of the "Corrected Item-Total Correlation" and " α if Item Deleted" statistics reveals that the removal of any individual item does not lead to a significant increase in the overall reliability coefficient. Consequently, there is no justification for deleting any items. In summary, the dataset exhibits high internal consistency and is deemed suitable for subsequent statistical analyses.

Table 5. Reliability Analysis

Dimension	Item	Correlation Between Adjusted Items and Totals	Cronbach's Alpha After Item Deletion	Total α Coefficient
SE	SE1	0.726	0.839	0.873
	SE2	0.730	0.837	
	SE3	0.720	0.841	
	SE4	0.737	0.834	
IE	IE1	0.762	0.884	0.905
	IE2	0.752	0.886	
	IE3	0.772	0.882	
	IE4	0.763	0.884	
	IE5	0.760	0.885	
FE	FE1	0.745	0.839	0.877
	FE2	0.728	0.846	
	FE3	0.750	0.837	
	FE4	0.719	0.849	
EE	EE1	0.766	0.836	0.880
	EE2	0.751	0.842	
	EE3	0.737	0.847	
	EE4	0.706	0.859	
TE	TE1	0.721	0.777	0.846
	TE2	0.721	0.777	
	TE3	0.695	0.802	

QoE	QoE1	0.686	0.782	0.835
	QoE2	0.685	0.783	
	QoE3	0.719	0.750	
US	US1	0.713	0.785	0.846
	US2	0.703	0.796	
	US3	0.724	0.776	
CUI	CUI1	0.686	0.796	0.840
	CUI2	0.719	0.765	
	CUI3	0.709	0.774	
AITS	AITS1	0.753	0.805	0.866
	AITS2	0.729	0.828	
	AITS3	0.755	0.804	

4.4 Validity Analysis

4.4.1 KMO and Bartlett's Test

Validity refers to the extent to which a test or scale can accurately measure the intended psychological and behavioral characteristics—that is, the accuracy and reliability of the test results. Generally, the lower the significance level of Bartlett's sphericity test ($P < 0.05$), the more likely it is that meaningful relationships exist among the original variables. The KMO value is used to compare simple and partial correlation coefficients among test items and ranges from 0 to 1. The criteria for suitability for factor analysis are as follows: greater than 0.9 indicates very suitable; 0.7–0.9 indicates suitable; 0.6–0.7 indicates somewhat suitable; 0.6–0.5 indicates not very suitable; and below 0.5 indicates unsuitable. The Bartlett's sphericity test is used to determine whether the correlation coefficients between items are significant. If the significance level is less than 0.05, it indicates that each item is suitable for factor analysis.

As shown in Table 6, the KMO value is 0.861, which is greater than 0.8; therefore, the research data are highly suitable for factor extraction (which indirectly reflects good validity). The KMO and Bartlett tests were computed on the 32 core measurement items only; the six demographic and screening questions were not included.

Table 6. KMO and Bartlett's Test

KMO Sampling Adequacy Measure		0.861
Bartlett Sphericity Test	Approximate Chi-Square	3,956.72
	Degrees of Freedom	496.000
	Significance	0.000

4.4.2. Exploratory Factor Analysis

This study conducted an exploratory factor analysis (EFA) using half of the sample ($n = 268$), employing principal axis factoring with Promax (oblique) rotation. An oblique rotation was selected because the substantive constructs are theoretically related and their observed correlations are moderate to high (roughly 0.5–0.75; see Table 10), making an orthogonal Varimax solution inappropriate. Factors with eigenvalues greater than 1 were retained. The nine factors jointly accounted for 76.00% of the variance before rotation (72.56% after Promax rotation), and every item had a primary pattern loading above 0.69, with no absolute cross-loading exceeding 0.28. Therefore, as shown in Table 7, nine common factors were extracted, consistent with the pre-specified dimensional structure and demonstrating good validity.

To address potential common method variance (CMV), Harman's single-factor test was conducted on all 32 core measurement items (excluding the six demographic and screening questions). The results revealed nine factors with eigenvalues greater than 1, and the first factor explained only 35.946% of the variance, which is below the 50% threshold. This indicates that CMV is not a serious concern in this study.

Table 7. Rotated Factor Pattern Matrix (Promax)

Item	SE	IE	FE	EE	TE	QoE	US	CUI	AITs
SE1	0.734	-0.080	0.051	-0.026	0.040	-0.003	-0.090	0.278	-0.048
SE2	0.931	0.085	-0.104	0.093	-0.052	0.025	0.102	-0.270	0.027
SE3	0.828	0.010	-0.024	0.007	0.047	-0.092	-0.062	0.141	0.036
SE4	0.842	-0.034	0.078	-0.064	0.006	0.082	0.018	-0.012	-0.032
IE1	-0.074	0.832	-0.075	0.081	0.129	-0.002	0.000	-0.013	0.009
IE2	-0.020	0.837	0.058	-0.036	-0.040	-0.085	0.034	0.116	-0.038
IE3	0.089	0.872	-0.127	-0.013	0.086	0.037	0.008	-0.089	-0.021
IE4	-0.006	0.785	0.152	-0.016	-0.141	0.121	-0.039	0.009	-0.003
IE5	0.013	0.795	0.058	-0.032	0.022	0.002	-0.072	0.013	0.074
FE1	-0.154	0.003	0.854	0.009	0.065	-0.031	0.077	-0.031	0.007
FE2	-0.053	0.010	0.857	0.012	-0.031	-0.027	0.012	0.078	0.035
FE3	0.114	0.086	0.697	0.026	0.003	0.029	-0.020	-0.006	-0.048
FE4	0.120	-0.027	0.861	-0.018	0.000	-0.014	0.016	-0.091	0.007
EE1	-0.046	0.084	-0.017	0.831	-0.005	-0.028	0.028	0.043	0.026
EE2	0.051	0.100	0.018	0.728	-0.177	-0.066	0.214	0.004	-0.044
EE3	0.002	-0.023	-0.070	0.832	0.021	-0.012	0.051	0.069	-0.011
EE4	0.034	-0.169	0.108	0.801	0.159	0.116	-0.220	-0.039	0.060
TE1	-0.045	-0.042	0.023	0.001	0.876	0.119	0.114	-0.147	-0.017
TE2	0.077	0.064	-0.006	0.094	0.774	-0.070	0.028	0.007	-0.095
TE3	0.015	0.074	0.035	-0.082	0.743	-0.070	-0.029	0.142	0.094
QoE1	0.008	-0.019	-0.113	-0.021	-0.024	0.881	0.100	0.040	0.083
QoE2	0.012	0.059	0.036	0.047	0.030	0.904	-0.098	-0.042	-0.110
QoE3	-0.010	0.042	0.042	-0.018	-0.002	0.769	0.037	0.126	0.025
US1	0.076	-0.068	0.104	0.013	0.013	0.078	0.749	0.019	0.030
US2	-0.065	-0.003	0.032	0.074	0.049	-0.050	0.904	-0.016	-0.068
US3	0.032	-0.013	0.008	-0.067	0.075	0.038	0.773	0.107	0.064
CUI1	-0.036	0.032	-0.046	0.111	-0.047	0.054	-0.045	0.904	-0.049
CUI2	-0.026	-0.013	-0.001	-0.076	0.018	0.037	0.128	0.853	-0.019
CUI3	0.095	0.033	-0.004	0.093	-0.006	-0.011	0.007	0.735	0.055
AITs1	0.219	0.019	-0.048	-0.021	-0.019	-0.049	0.062	-0.076	0.877
AITs2	-0.021	0.020	0.111	0.002	-0.052	0.036	-0.044	-0.051	0.869
AITs3	-0.188	-0.020	-0.054	0.052	0.051	-0.000	-0.019	0.093	0.877

Note: Principal axis factoring with Promax oblique rotation; entries are standardized pattern coefficients from the EFA half-sample ($n = 268$). The lowest primary loading is 0.697 (FE3), and the largest absolute cross-loading is 0.278.

4.4.3. Confirmatory Factor Analysis

Confirmatory factor analysis (CFA) was conducted on the second, independent random half of the sample ($n = 268$) to cross-validate the factor structure obtained in the EFA. As shown in Table 8, the standardized CFA loadings ranged from 0.68 to 0.88, and the measurement model fit the holdout data well: $\chi^2 = 591.53$, $df = 428$, $\chi^2/df = 1.38$, CFI = 0.965, TLI =

0.960, RMSEA = 0.038, and SRMR = 0.034. These results confirm the nine-factor measurement structure on an independent sample.

Table 8. Confirmatory Factor Analysis (Holdout Sample, n = 268)

Construct	Item	Standardized loading
SE	SE1	0.809
	SE2	0.758
	SE3	0.812
	SE4	0.746
IE	IE1	0.810
	IE2	0.800
	IE3	0.809
	IE4	0.814
	IE5	0.782
FE	FE1	0.757
	FE2	0.788
	FE3	0.751
	FE4	0.779
EE	EE1	0.748
	EE2	0.684
	EE3	0.792
	EE4	0.749
TE	TE1	0.699
	TE2	0.760
	TE3	0.724
QoE	QoE1	0.800
	QoE2	0.758
	QoE3	0.828
US	US1	0.818
	US2	0.753
	US3	0.819
CUI	CUI1	0.828
	CUI2	0.766
	CUI3	0.878
AITS	AITS1	0.773
	AITS2	0.849
	AITS3	0.818
Model fit	χ^2 (df = 428)	591.527
	χ^2/df	1.382
	CFI	0.965
	TLI	0.960
	RMSEA	0.038
	SRMR	0.034

Note: CFA loadings are standardized maximum-likelihood estimates from the independent holdout half-sample (n = 268); model fit indices are reported in the last six rows.

4.5. Measurement Model

4.5.1. Convergent Validity

When conducting a convergent validity analysis, it is generally necessary to examine the following indicators: Average Variance Extracted (AVE), Composite Reliability (CR), and factor loading coefficients. Typically, if all AVE values are greater than 0.5 and all CR values are higher than 0.7, the convergent validity is considered high. As shown in Table 9, the AVE values for all factors are greater than 0.5, and the CR values are all higher than 0.7, indicating that the data in this analysis possess good convergent validity.

Table 9. AVE and CR Values

Dimension	CR	AVE
SE	0.873	0.633
IE	0.905	0.656
FE	0.878	0.642
EE	0.880	0.648
TE	0.846	0.647
QoE	0.836	0.630
US	0.846	0.648
CUI	0.841	0.638
AITS	0.871	0.693

4.5.2 Discriminant Validity

Discriminant validity was assessed using the Fornell–Larcker criterion. Specifically, if the square root of the Average Variance Extracted (AVE) for each dimension (values on the diagonal in bold in Table 10) is greater than the absolute value of the correlation between that construct and any other construct (off-diagonal values), discriminant validity is considered good. As shown in Table 10, all diagonal values (square roots of AVE) are consistently greater than the corresponding off-diagonal values in their rows and columns, indicating satisfactory discriminant validity. The off-diagonal entries in Table 10 are Pearson correlations between construct composite scores (the unweighted mean of each construct's standardized items). Because these composite correlations contain measurement error, they are attenuated relative to the disattenuated latent correlations and structural paths; for example, the composite QoE–US correlation (0.638) is lower than the standardized latent QoE→US path (0.79), which is the expected direction and magnitude of attenuation (the disattenuated correlation is approximately 0.76–0.79).

Table 10. Discriminant Validity

	SE	IE	FE	EE	TE	QoE	US	CUI	AITs
SE	0.796								
IE	0.538***	0.810							
FE	0.604***	0.645***	0.801						
EE	0.561***	0.535***	0.631***	0.805					
TE	0.577***	0.568***	0.633***	0.554***	0.804				
QoE	0.570***	0.600***	0.638***	0.559***	0.606***	0.793			
US	0.591***	0.626***	0.685***	0.597***	0.633***	0.638***	0.805		
CUI	0.674***	0.692***	0.719***	0.708***	0.688***	0.671***	0.757***	0.799	
AITs	0.229***	0.359***	0.293***	0.255***	0.290***	0.175***	0.359***	0.237***	0.828

4.6 Structural Model Analysis

4.6.1 Model Fit Indices

This study assessed the model fit, as a good fit indicates that the model closely approximates the data. Multiple fit indices— χ^2 and degrees of freedom, CMIN/DF, GFI, NFI, IFI, TLI, CFI, RMSEA, and SRMR [90]—were selected to evaluate the structural model. The structural model was estimated on the full sample ($n = 536$) using maximum likelihood. As shown in Table 11, the structural model fit the data well ($\chi^2 = 837.90$, $df = 441$, CMIN/DF = 1.9, SRMR = 0.058), confirming that the theoretical framework underlying the research hypotheses is consistent with the findings.

Table 11. Model Fit Metrics

Metrics	Measurement Criteria	Data from this study	Evaluation Results
CMIN/DF	<3	1.9	Yes
GFI	>0.9	0.917	Yes
NFI	>0.9	0.932	Yes
IFI	>0.9	0.967	Yes
TLI	>0.9	0.962	Yes
CFI	>0.9	0.966	Yes
RMSEA	<0.08	0.041	Yes
χ^2	—	837.90	—
df	—	441	—
SRMR	<0.08	0.058	Yes

4.6.2 Structural Equation Path Analysis

As shown in Table 12, older users' SE toward mobile health apps ($b = 0.147$, $P < 0.001$) has a positive effect on QoE, indicating that Hypothesis H1 is supported; older users' IE toward mobile health apps ($b = 0.202$, $P < 0.001$) indicates that Hypothesis H2 is supported; Elderly users' FE ($b = 0.187$, $p < 0.001$) regarding mobile health applications has a positive effect on QoE, meaning Hypothesis H3 is supported; elderly users' EE ($b = 0.151$, $p < 0.001$) regarding mobile health applications has a positive effect on QoE, meaning Hypothesis H4 is supported; Older users' TE ($b = 0.198$, $p < 0.001$) regarding mHealth apps has a positive effect on QoE, meaning Hypothesis H5 is supported; QoE ($b = 0.936$, $p < 0.001$) has a positive effect on US, meaning Hypothesis H6 is supported; QoE ($b = 0.815$, $p < 0.001$) has a positive effect on user CUI, meaning Hypothesis H7 is supported; US ($b = 0.198$, $p = 0.004$) has a positive effect on user CUI, meaning Hypothesis H8 is supported. Based on the results of the structural equation path analysis, it can be seen that the relationships among the variables are all significant.

Table 12. Path Analysis

Path	Unstd	S.E.	STD	C.R.	p
SE ---> QoE	0.147	0.037	0.18	4.015	***
IE ---> QoE	0.202	0.038	0.24	5.292	***
FE ---> QoE	0.187	0.043	0.24	4.397	***
EE ---> QoE	0.151	0.036	0.19	4.212	***
TE ---> QoE	0.198	0.041	0.24	4.858	***
QoE ---> US	0.936	0.068	0.79	13.842	***
QoE ---> CUI	0.815	0.093	0.70	8.809	***
US ---> CUI	0.198	0.069	0.20	2.856	0.004

Note: *** $p < 0.001$; p -values are two-tailed. Unstd = unstandardized estimate; S.E. = standard error; STD = standardized estimate; C.R. = critical ratio.

Pairwise critical-ratio tests for the five experience-dimension paths, computed as $z = (b_i - b_j) / \sqrt{(SE_i^2 + SE_j^2)}$, were all non-significant (all $|z| < 1.96$, $p > 0.05$): SE-IE = -1.04, SE-FE = -0.71, SE-EE = -0.08, SE-TE = -0.92, IE-FE = 0.26, IE-EE = 0.97, IE-TE = 0.07, FE-EE = 0.64, FE-TE = -0.19, and EE-TE = -0.86. Thus, no two experience-dimension coefficients differ statistically from one another.

4.6.3 Analysis of Moderating Effects

To test the moderating roles of AITS, latent interaction terms were constructed using the product-indicator approach. The items of each focal dimension and of AITS were mean-centered to reduce non-essential multicollinearity, and pairwise product indicators were formed; the latent interaction factors were then estimated in AMOS 17.0 with maximum likelihood, with the measurement portions of the focal and moderator factors fixed to their confirmatory factor values. Both unstandardized coefficients (B) and standardized moderation coefficients ($STD-\beta$) were estimated and are presented in Table 13. The results show that AITS exerts a significant positive moderating effect on the

SE-QoE relationship ($B = 0.102$, $STD-\beta = 0.112$, $t = 4.179$, $p < 0.001$), and similarly on the IE-QoE ($B = 0.054$, $STD-\beta = 0.061$, $t = 2.265$, $p = 0.024$), FE-QoE ($B = 0.079$, $STD-\beta = 0.086$, $t = 3.576$, $p < 0.001$), and EE-QoE ($B = 0.073$, $STD-\beta = 0.078$, $t = 2.999$, $p = 0.003$) relationships. Nevertheless, the moderating effect on the TE-QoE relationship is not statistically significant ($B = 0.041$, $STD-\beta = 0.045$, $t = 1.749$, $p = 0.081$; 95% CI for $B = -0.004$ to 0.086). Accordingly, H9a, H9b, H9c and H9d are supported, whereas H9e is rejected. A post-hoc power analysis indicated that the TE interaction test had approximately 42% power to detect the observed standardized effect at $\alpha = 0.05$, so the non-significant result should be interpreted as inconclusive rather than as evidence of no effect. The magnitude of the standardized coefficients indicates that AITS produces the strongest amplification on the sensory pathway, followed by functional and emotional experience, with a comparatively weak effect on interaction. Simple-slope plots for the four significant moderations (at $AITS = M \pm 1SD$) are presented in Figure 3.

Table 13. Analysis of the Moderating Effects of AITS

Moderating Path	Unstandardized B	S.E.	STD- β (moderation)	t	Significance
AITS $\times$ SE $\rightarrow$ QoE	0.102	0.024	0.112	4.179	<0.001**
AITS $\times$ IE $\rightarrow$ QoE	0.054	0.024	0.061	2.265	0.024*
AITS $\times$ FE $\rightarrow$ QoE	0.079	0.022	0.086	3.576	<0.001**
AITS $\times$ EE $\rightarrow$ QoE	0.073	0.024	0.078	2.999	0.003**
AITS $\times$ TE $\rightarrow$ QoE	0.041	0.023	0.045	1.749	0.081

Note: * $p < 0.05$; ** $p < 0.01$ (two-tailed); p -values below 0.001 are reported as <0.001. B = unstandardized coefficient; $STD-\beta$ = standardized moderation coefficient. The AITS $\times$ TE interaction did not reach the $p < 0.05$ threshold.

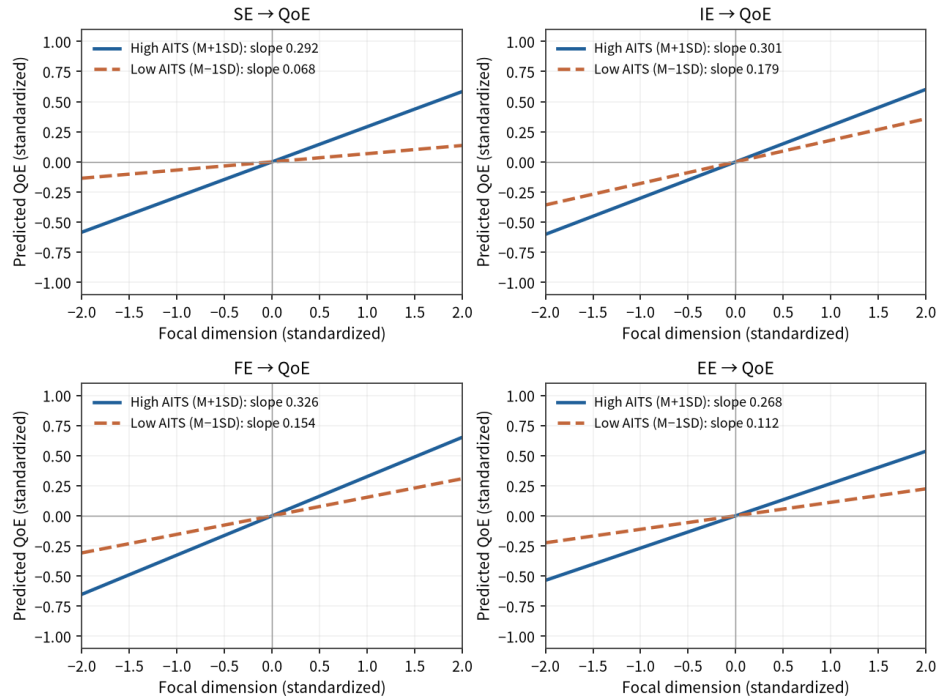


Figure 3. Simple-Slope Plots for the Significant Moderating Effects of AITS

4.6.4 Analysis of Mediation (Indirect) Effects

To test the mediating role of QoE, indirect effects were estimated in AMOS 17.0 with 5,000 bootstrap resamples using bias-corrected 95% confidence intervals [88]. As reported in Table 14, the total indirect effects of each experience dimension on CUI through QoE were all positive and significant (e.g., IE: 0.206, 95% CI 0.136–0.276; FE: 0.206, 95% CI 0.122–0.289), and the indirect effects of the dimensions on US via QoE likewise excluded zero. The serial indirect effect of QoE on CUI through US was 0.158 (95% CI 0.052–0.264). These results empirically establish QoE's mediating role in translating experience dimensions into satisfaction and continued-use intention.

Table 14. Bootstrap Indirect Effects (Bias-Corrected 95% CI)

Path	Indirect effect	BC 95% CI
SE → QoE → US	0.142	[0.072, 0.212]
IE → QoE → US	0.190	[0.123, 0.256]
FE → QoE → US	0.190	[0.108, 0.271]
EE → QoE → US	0.150	[0.074, 0.226]
TE → QoE → US	0.190	[0.112, 0.268]
SE → QoE → CUI (total indirect)	0.154	[0.079, 0.229]

IE → QoE → CUI (total indirect)	0.206	[0.136, 0.276]
FE → QoE → CUI (total indirect)	0.206	[0.122, 0.289]
EE → QoE → CUI (total indirect)	0.163	[0.083, 0.243]
TE → QoE → CUI (total indirect)	0.206	[0.124, 0.288]
QoE → US → CUI (serial)	0.158	[0.052, 0.264]

4.7 Hypothesis Interpretation

Based on the results of the structural equation path analysis and moderation tests, most hypothesized relationships among the variables were statistically significant. Specifically, the moderating effect of AITS on the TE–QoE relationship (H9e) was not significant ($p = 0.081 > 0.05$). Therefore, Hypotheses H1 through H8 and H9a through H9d are supported, while H9e is not supported. The order of influence of the significant factors is as follows (as shown in Figure 4): "QoE → US, 0.79" > "QoE → CUI, 0.70" > "IE → QoE, 0.24" = "FE → QoE, 0.24" = "TE → QoE, 0.24" > "US → CUI, 0.20" > "EE → QoE, 0.19" > "SE → QoE, 0.18." This ordering is based on point estimates; the pairwise critical-ratio tests reported in the note to Table 12 indicate that the differences among the five experience-dimension coefficients (0.18–0.24) are not statistically significant, so the relative ordering of these dimensions should be interpreted as exploratory.

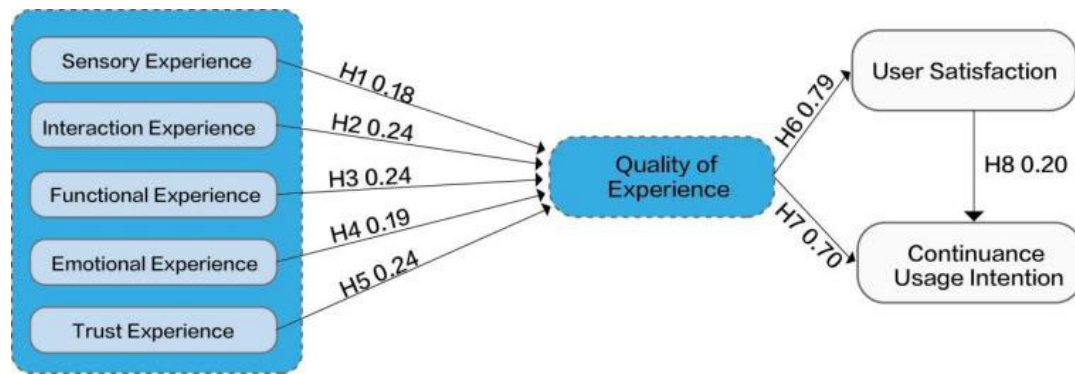


Figure 4. Regression Coefficients of the SEM Model

5. Results and Discussion

This study empirically examines the formation mechanisms of CUI among older adult users of AI-enabled mHealth services. It aims to provide an in-depth, multi-level interpretation of the findings, with a particular focus on the moderating role of AI technology and the unique characteristics of the elderly demographic, thereby enhancing the theoretical depth and practical relevance of this research.

5.1 Heterogeneity in the Influence Strength Across Experience Dimensions

The results confirm that all five experience dimensions significantly and positively impact QoE. In point-estimate terms, IE, FE, and TE (all $\beta=0.24$) are the highest of the five antecedents, whereas SE is the lowest ($\beta=0.18$); however, pairwise critical-ratio tests indicate that the differences among the five coefficients (0.18–0.24) are not statistically significant (see Section 6.2.1), so these differences should be read as exploratory rather than definitive. The overall pattern nonetheless underscores that, for elderly users, the core value proposition of AI healthcare services remains rooted in practicality and usability. AI technology should not be viewed merely as a vehicle for technological ostentation; its ultimate utility lies in addressing essential health management needs with greater accuracy and convenience. For instance, AI medical consultations capable of understanding local dialects and one-click appointment booking functions offer far greater value than visually appealing but impractical interface effects.

The high point estimate for TE also highlights the unique stakes of the healthcare sector. Privacy and reliability carry particular weight in this high-stakes context. The present study has no comparison group of general users and therefore cannot establish that older users worry “significantly more”; nevertheless, the strong TE coefficient indicates that if the “algorithmic black box” issue and data security risks are not effectively mitigated, any functional optimizations may ultimately be offset by a trust deficit.

Conversely, a somewhat counterintuitive finding is that EE ($\beta=0.19$), despite having the lowest mean score ($M = 4.71$; Table 4), still maintains a significant influence coefficient. This reveals a core contradiction and a substantial opportunity within current AI healthcare services: while older users possess a strong latent demand for emotional support, existing technologies fail to meet this need effectively.

This gap exists not because users lack the need, but because current AI-driven emotional interactions are largely predicated on pattern recognition rather than authentic understanding, resulting in a perceived superficial experience. On the engineering side, the counterpart metrics for EE—such as emotion-recognition accuracy, conversational empathy/agreement scores, and the alignment of response sentiment with the user's affect—were not measured in this study and should be included in future systems evaluations. EE currently represents the ‘weak link’ in service offerings; developers who achieve breakthroughs in this dimension through human-centered dialogue design or more compassionate interaction flows stand to gain a significant competitive advantage.

5.2 The Driving Mechanism of Continued Usage Intention

Path analysis reveals a robust transmission mechanism: QoE $\rightarrow$ US $\rightarrow$ CUI. Notably, the influence of QoE on US ($\beta=0.79$) and its direct effect on CUI ($\beta=0.70$) far exceed the direct effect of US on CUI ($\beta=0.20$). This finding carries significant theoretical implications.

It suggests that for older users, decisions regarding continued use resemble a rational, cognition-based judgment process rather than one driven solely by transient affect. Users rely more heavily on a comprehensive assessment of the service's overall excellence and reliability (i.e., QoE) than on momentary satisfaction alone. This challenges the traditional “US-centric” paradigm in user behavior research and illustrates that in the high-stakes health domain—particularly among older adults who tend to adopt a cautious cognitive style—building a stable, high-quality service experience is paramount. Consequently, enterprises should reallocate

resources, shifting their strategic focus from short-term satisfaction enhancement to the long-term, systematic cultivation of QoE.

5.3 The Moderating Role of AITS

The core theoretical innovation of this research is the incorporation of AITS as a moderating variable, which unpacks the heterogeneous boundary conditions of AI-enabled mHealth services and yields nuanced insights into the practical value of AI technologies for older-adult end users.

The empirical results demonstrate that AITS exerts statistically significant positive moderating effects on the predictive paths from SE, IE, FE, and EE to QoE. This evidence indicates that AI technology functions as an experience “amplifier”: it leverages inherent technical strengths, including natural-language-processing-driven conversational fluency and big-data-powered functional precision, to strengthen how positive user-experience perceptions translate into elevated overall experience quality. For instance, a well-refined sensory-oriented interface paired with an AI voice assistant can generate synergistic benefits (a “1+1>2” effect) and substantially mitigate the operational burdens faced by older adult users.

Nevertheless, the non-significant moderating influence of AITS on the TE–QoE linkage ($B = 0.041$, $p = 0.081$, 95% CI -0.004 to 0.086) is suggestive but not conclusive: the data neither establish nor rule out AI moderation of trust. Contemporary AI medical chatbots are frequently constrained by limited explainability: their internal reasoning remains opaque, forming the well-known “black-box” problem, and they possess insufficient fault tolerance when processing high-stakes medical inquiries. To mitigate such concerns, developers should prioritize three engineering-oriented design interventions. First, deploy explainable AI (XAI) solutions to improve algorithmic transparency, for example by delivering plain-text justifications behind

diagnostic recommendations and visualizing model confidence scores for end-users. Second, establish reliable human–AI fallback workflows, such as one-click human-agent transfer and predefined escalation rules for high-risk decision-making. Third, embed privacy-by-design principles throughout development, including end-to-end encryption and granular, explicit consent mechanisms. Proposed measurable engineering targets for these interventions are listed in Table 15; they should be treated as targets for future implementation and validation rather than as outcomes demonstrated by the present data. Collectively, these observations suggest that older adults’ trust is unlikely to be fostered by AI-centric features alone and may also rest upon non-technical safeguards such as institutional guarantees, reputable service providers, and human-supervision arrangements. A hybrid-intelligence design paradigm that combines the operational efficiency of AI with human-centered reliability therefore represents a prudent strategy for cultivating long-term trust.

Table 15. Proposed Engineering Targets for Age-Friendly AI mHealth Services

Dimension	Engineering target	Proposed threshold
SE	Interface rendering / response latency	≤ 0.5 s (≤ 1.0 s voice round-trip)
SE	Color contrast; font scaling	$\geq 4.5:1$ (WCAG 2.2 AA); scalable to 200%
IE	Dialect speech recognition word error rate	$\leq 15\%$ (accuracy $\geq 90\%$)
IE	Time to human handover	≤ 30 s, one-click transfer
FE	Diagnostic sensitivity / specificity	≥ 0.90
FE	System uptime	$\geq 99.9\%$
TE	XAI rationale and confidence display	For every high-risk output

TE	Data protection	Differential privacy; end-to-end encryption
EE	Emotion recognition / empathy accuracy	To be benchmarked in future work
EE	Warm, sentiment-aligned responses	Qualitative A/B evaluation

6. Conclusion

6.1 Theoretical Contributions

6.1.1 Revision and Reconstruction of the Classical Model for an Aging Context

This study provides a contextual extension and age-appropriate revision of the classical IS continuance model. While traditional models emphasize a linear path of "expectation confirmation → satisfaction → continuance," this study positions the multidimensional construct of "experience" at the forefront. The central mediating role of QoE is supported by the bootstrap indirect-effect tests in Table 14. Age-related conditions—declining sensory-motor function, cognitive changes, and technology anxiety—were treated as background considerations that informed the design of the items and the sampling framework; they were not directly measured as variables and should therefore not be understood as empirically tested components of the model.

Based on these findings, this paper preliminarily proposes an "Adapted Age-Friendly Technology Acceptance Framework," positing that for older users with high health engagement, the mechanism of technology acceptance shifts from reliance on abstract cognitions (e.g., perceived usefulness) toward dependence on a concrete, multidimensional experience quality. It should be acknowledged that the integration of ECM-ISC is partial: the model retains the satisfaction → continuance core and substitutes QoE and the five experience dimensions for confirmation and perceived usefulness, but confirmation and perceived usefulness were not directly measured. The framework is also proposed based on cross-sectional intention data; longitudinal validation of actual usage behavior and health outcomes is needed before it can be established as a definitive "model."

6.1.2 Revealing the Boundaries of AI Technology's Moderating Effects

Empirical analysis reveals the following structural relationships: (1) SE ($\beta=0.18$), IE ($\beta=0.24$), FE ($\beta=0.24$), EE ($\beta=0.19$), and TE ($\beta=0.24$) all significantly and positively affect QoE, with IE and FE exhibiting the most pronounced effects; (2) QoE not only significantly enhances US ($\beta=0.79$) but also directly promotes CUI ($\beta=0.70$), establishing it as a pivotal driver of retention; (3) US has a direct positive effect on CUI ($\beta=0.20$), though its effect is weaker than that of QoE; (4) AITS exerts a significant positive moderating effect on the influence of SE, IE, FE, and EE on QoE, but its moderating effect on TE is statistically non-significant.

This delineates the mechanism of AI technology's moderating effects. The finding that AITS significantly moderates the sensory, interactive, functional, and emotional dimensions (e.g., standardized moderation coefficient for SE = 0.112) but not the technological/trust dimension, refines the boundaries of AI's role. It provides a theoretical basis for understanding the differentiated pathways through which technology enhances user experience in high-stakes domains.

6.2 Practical Implications

Based on the foregoing discussion, the practical implications of this study suggest a strategic upgrade beyond generic design principles:

6.2.1 Experience Prioritization and Tiered Design Strategy

Resource investment should follow a priority hierarchy. Functionality, trust, and interaction form the first tier ($\beta = 0.24$): priority must be given to the accuracy, reliability, and data security of core functions, which form the cornerstone of trust, and AI should be leveraged to optimize the naturalness and fault tolerance of interactions. Sensory and emotional experience form the lower tier; in point-estimate terms EE ($\beta = 0.19$) is slightly above SE ($\beta = 0.18$), but pairwise critical-ratio tests indicate that the differences among the five coefficients (0.18–0.24) are not statistically significant, so this ordering should be treated as exploratory. Given that EE also has the lowest mean, a pragmatic, complementary strategy for the emotional dimension is warranted, while avoiding overemphasis on human–AI emotional interaction, which remains technologically immature.

6.2.2 Trust-Building Strategies: Implementing "Trustworthy AI"

"Service providers should launch a transparency initiative. For instance, when AI provides recommendations, it should include plain-language explanations (e.g., "Based on your description of 'dull pain after meals,' there is a high likelihood this relates to gastritis. Consider the following dietary precautions"). Furthermore, establishing a clear "One-Click Transfer to Human Agent" channel is crucial to define the limits of AI capability. To mitigate uncertainty from algorithmic black boxes, systems should allow elderly users (with informed consent) to authorize children or guardians to receive critical health alerts and AI decision logs, thereby constructing a three-tiered trust defense system comprising "AI Execution + Human Fallback + Family Supervision."

6.2.3 Emotional Compensation Strategies

Until AI achieves genuine empathy, emotional compensation should be realized through design details. For example, employing a warm, encouraging tone and providing positive visual or textual feedback (e.g., "Well done! Your health report has been successfully generated") after users complete complex operations can help offset the perceived coldness of technology.

6.3 Limitations and Future Directions

Certain limitations of this study point to avenues for future research:

First, Sample Scope and Regional Generalizability. The data collected were limited to older adults in specific regions of East China. Given the significant disparities in economic development and digital infrastructure across mainland China, usage patterns may vary substantially among older populations in different regions. Future research should expand the sample to include diverse geographical areas to enhance the generalizability of the findings.

Second, Heterogeneity of Application Types. Given the diversity of AI-enabled mHealth applications (e.g., chronic disease management vs. teleconsultation), variations in usage levels may exist across different software categories. Future studies should focus on distinct types of AI health applications to develop corresponding theoretical models tailored to specific service contexts.

Third, Individual Health and Cognitive Differences. Older users face varying degrees of physical illness and cognitive decline, which may further affect their interaction with information technology.

Therefore, future research should explore the moderating roles of physical health status and cognitive ability, examining how these individual difference variables impact the sustained use of AI-enabled mobile health applications [89].

Finally, it is important to note that this study measures continued usage intention rather than actual longitudinal continuance behavior or health outcomes. The claims regarding long-term engagement should therefore be interpreted with caution, and future research should incorporate objective usage logs and longitudinal designs to validate the proposed framework.

Data Availability: The data supporting the findings of this study are available from the corresponding author upon reasonable request, subject to the ethical and privacy restrictions that apply to data collected from older participants.

Acknowledgements

This research was supported by the Humanities and Social Sciences Fund of the Ministry of Education of China (Grant No. 25YJCZH266), the Talent Introduction and Start-up Research Fund of Nanjing Institute of Technology (Grant No. YKJ202457), the National Project on Application and Innovation of AI General Education Courses in Higher Vocational Colleges (Grant No. KT2510077), the Humanities and Social Sciences Research Program of the Korean Academy of Humanities and Social Sciences

(Grant No. 2025KIHSSB10), the Higher Education Research Project of Nanjing Institute of Technology (Research on High-Quality Training Models for Design Education under the Double-Edged Sword Effect of Artificial Intelligence), the Research Project of the China Arts and Crafts Society (Grant No. CNACS20225-B-I-31), and the Research Project of the China Ethnic and Folk Arts and Crafts Artists Association (Grant No. YBKT2025012).

References

- [1] National Bureau of Statistics of China. 2025 national 1% population sampling survey: Main data bulletin. Available at: https://www.stats.gov.cn/sj/zxfb/202605/t20260522_1963788.html (accessed May 22, 2026).
- [2] Liu WL, Yang Y, Li ZH, Lin B, Hao XN, Zhang WJ, Xie LL. Strategies for China's response to population aging under the new situation. *Strategic Study of CAE*. 2024;26(6):19–31.
- [3] Wang LM, Chen ZH, Zhang M, Zhao ZP, Huang ZJ, Zhang X, Zhou MG. Chronic disease prevalence and disease burden among the elderly in China. *Chinese Journal of Epidemiology*. 2019;40(3):277–283.
- [4] Wang Y, Wang X, Ye X. Inequalities in unmet needs for healthcare services among middle-aged and older adults in China. *Journal of Aging & Social Policy*. 2026;38(2):259–276. <https://doi.org/10.1080/08959420.2024.2422674>
- [5] Liu Y, Qin Y, Chen Q, Liu Y, Jia A, Liu M, Ji N, Zhang Z, Shao R, Bai Y. Effectiveness of mobile health interventions on physical activity management in adults with hypertension: A systematic review and meta-analysis. *International Journal of Nursing Studies*. 2025;172:105225.
- [6] Lu J, Zhao H, Ji Z, Dong Y. Effect of patient satisfaction on the utilization of mHealth services by patients with chronic disease. *Digital Health*. 2025;11.
- [7] Hu E, Wang H, Yao J, Liao M, Chen W. Impact of digital health interventions on quality of life and mental health in older adults with chronic diseases: A systematic review and meta-analysis. *Frontiers in Aging*. 2026;6:1737277. <https://doi.org/10.3389/fragi.2025.1737277>
- [8] Long JH, Yi MH. Research on the mechanism of digital technology in alleviating urban-rural elderly care service disparities in China. *Agricultural Archaeology*. 2025;(3):22–29.
- [9] Feng G, Weng F, Lu W, Xu L, Zhu W, Tan M, Weng P. Artificial intelligence in chronic disease management for aging populations: A systematic review of machine learning and NLP applications. *International Journal of General Medicine*. 2025;18:3105–3115. <https://doi.org/10.2147/IJGM.S516247>
- [10] China Internet Network Information Center. The 55th statistical report on China's internet development. Available at: <https://www.cnnic.com.cn/IDR/ReportDownloads/202505/P020250514564119130448.pdf> (accessed May 28, 2026).
- [11] Big Data Network. Q1 2021 China mobile health product market monitoring report. Baijiahao [online]. Available at: <https://baijiahao.baidu.com/s?id=1698393186006070449> (accessed 2021).
- [12] Anisha SA, Sen A, Ahmad B, et al. Exploring acceptance of digital health technologies for managing non-communicable diseases among older adults: A systematic scoping review. *Journal of Medical Systems*. 2025;49:35. <https://doi.org/10.1007/s10916-025-02166-3>
- [13] Wong AKC, Lee JH, Zhao Y, Lu Q, Yang S, Hui VCC. Exploring older adults' perspectives and acceptance of AI-driven health technologies: Qualitative study. *JMIR Aging*. 2025;8:e66778.

- [14] Amouzadeh E, Dianat I, Faradmal J, et al. Optimizing mobile app design for older adults: Systematic review of age-friendly design. *Aging Clinical and Experimental Research*. 2025;37:248. <https://doi.org/10.1007/s40520-025-03157-7>
- [15] Johnson D, Goodman R, Patrinely J, Stone C, Zimmerman E, Donald R, et al. Assessing the accuracy and reliability of AI-generated medical responses: An evaluation of the Chat-GPT model. *Research Square*. 2023;rs.3.
- [16] Lee D, Yoon SN. Application of artificial intelligence-based technologies in the healthcare industry: Opportunities and challenges. *International Journal of Environmental Research and Public Health*. 2021;18(1):271.
- [17] Hindelang M, Sitaru S, Zink A. Transforming health care through chatbots for medical history-taking and future directions: Comprehensive systematic review. *JMIR Medical Informatics*. 2024;12:e56628.
- [18] Amiri Z, Taghavirashidizadeh A, Khorrami P. AI-driven decision-making in healthcare information systems: A comprehensive review. *Journal of Systems and Software*. 2025;112470.
- [19] Chow J, Wong V, Li K. Generative pre-trained transformer-empowered healthcare conversations: Current trends, challenges, and future directions in large language model-enabled medical chatbots. *BioMedInformatics*. 2024;4(1):837–852.
- [20] Song I, Pendse S, Kumar N, Choudhury M. The typing cure: Experiences with large language model chatbots for mental health support. *Proceedings of the ACM on Human-Computer Interaction*. 2025;9(7).
- [21] Alon Y, Naimi E, Levin C, Videll H, Saban M. Leveraging natural language processing to elucidate real-world clinical decision-making paradigms: A proof of concept study. *Journal of Biomedical Informatics*. 2025;166.
- [22] Masood S, Bashrawi M, Khan M, Nazir A. Exploring ChatGPT applications in healthcare: A comprehensive overview. *The Visual Computer*. 2024;1–22.
- [23] He C, Deng Y. AI-assisted inclusive interaction design system for users with cognitive differences.
- [24] Salih H, Nsaif W, Al-Nuaimi B, Saleh H. Exploring the characteristics, technologies and challenges of healthcare chatbots. *Proceedings of the International Conference on Applied Innovations in IT*. 2025;13(2):155–165.
- [25] Ghorbani M, Ghobaei-Arani M. Large language models for mental health diagnosis and treatment: A survey. *Artificial Intelligence Review*. 2025;59(1):1–50.
- [26] Li H S, Gai J P, Bu L Y, et al. Enhancing universal access in elderly health information systems: the triumvirate of trust, adherence, and cognitive load with explainable AI. *Universal Access in the Information Society*, 2025, 24(4): 3449–3463.
- [27] Cozzolino A, Geiger S. Ecosystem disruption and regulatory positioning: Entry strategies of digital health startup orchestrators and complementors. *Research Policy*. 2024;53(2).
- [28] Weimar S, Martjan R, Terzidis O. Business venturing in regulated markets—Taxonomy and archetypes of digital health business models in the European Union: Mixed methods descriptive and exploratory study. *Journal of Medical Internet Research*. 2025;27:e65725.
- [29] Jogova M, Shaw J, Jamieson T. The regulatory challenge of mobile health: Lessons for Canada. *Healthcare Policy*. 2019;14(3):19.
- [30] Zhang Q, Liu Y. Smart user experience medical app interface design based on mobile devices. *Expert Systems*. 2022;39(5):e12808.
- [31] Zhang L, Zhan J, Wan V, Wang Y. Designing and evaluating online health consultation interfaces: A perspective of physician-patient power asymmetry. *IEEE Access*. 2024;12:124111–124127.
- [32] Kitsiou S, Gerber BS, Buchholz SW, Kansal MM, Sun J, Pressler SJ. Patient-centered mHealth intervention to improve self-care in patients with chronic heart failure: Phase 1 randomized controlled trial. *Journal of Medical Internet Research*. 2025;27:e55586.
- [33] Ruiz-Leon A, Casas R, Castro-Barquero S, Alfaro-González S, Radeva P, Sacanella E, et al. Efficacy of a mobile health-based behavioral treatment for lifestyle modification in type 2 diabetes self-management: Greenhabit randomized controlled trial. *Journal of Medical Internet Research*. 2025;27:e58319.
- [34] Candelaria D, Cacciata M, Serafica R, et al. Patient activation improves with a multi-component personalized mHealth intervention in older patients at risk of cardiovascular disease: a pilot randomized controlled trial. *European Journal of Cardiovascular Nursing*, 2025, 24(2): 316–322.
- [35] Wang X, Lee CF, Jiang J, Zhu X. Factors influencing the aged in the use of mobile healthcare applications: An empirical study in China. *Healthcare*. 2023;11(3):396. <https://doi.org/10.3390/healthcare11030396>
- [36] Wang X, Lee CF, Jiang J, Zhang G, Wei Z. Research on the factors affecting the adoption of smart aged-care products by the aged in China: Extension based on UTAUT model. *Behavioral Sciences*. 2023;13(3):277. <https://doi.org/10.3390/bs13030277>
- [37] Bhattacharjee A. Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly*. 2001;25(3):351–370.
- [38] Siller M, Woods JC. QoS arbitration for improving the QoE in multimedia transmission. In: 2003 International Conference on Visual Information Engineering (VIE 2003); 2003; IET. p. 238–241.
- [39] Akter S, D'Ambra J, Ray P. Service quality of mHealth platforms: Development and validation of a hierarchical model using PLS. *Electronic Markets*. 2020;20(3–4):209–227.
- [40] Hassenzahl M, Tractinsky N. User experience—A research agenda. *Behaviour & Information Technology*. 2006;25(2):91–97.
- [41] Hassenzahl M, Diefenbach S, Göritz A. Needs, affect, and interactive products—Facets of user experience. *Interacting with Computers*. 2010;22(5):353–362.
- [42] Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*. 1989;13(3):319–340.
- [43] DeLone WH, McLean ER. The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*. 2003;19(4):9–30.
- [44] Saariluoma P, Jokinen JP. Emotional dimensions of user experience: A user psychological analysis. *International Journal of Human-Computer Interaction*. 2014;30(4):303–320.
- [45] Ajith TN, S P, Mathew LS. Experience matters: Exploring the impact of user experience on stickiness and loyalty in OTT platforms. *International Journal of Human-Computer Interaction*. 2025;41(18):11363–11377.
- [46] Bélanger F, Crossler RE. Privacy in the digital age: A review of information privacy research in information systems. *MIS Quarterly*. 2011;35(4):1017–A36.
- [47] Song T. The impact of shape and decoration on user experience and visual attention in anthropomorphic robot design. *Journal of Eye Movement Research*. 2025;18(2):5.
- [48] Zhao WJ, Xie SM. Relationship among perceived value, usability (US) and behavioral intention in college students'

- mobile reading: Taking Chaoxing mobile reading APP platform as an example. *Library and Information Service*. 2019;63(3):98–107 [in Chinese]. <https://doi.org/10.13266/j.issn.0252-3116.2019.03.013>
- [49] Poushneh A, Vasquez-Parraga AZ. Discernible impact of augmented reality on retail customer's experience, satisfaction and willingness to buy. *Journal of Retailing and Consumer Services*. 2017;34:229–234.
- [50] Cao JY, Zhang M, He HY, Liu HT. Research on the influence of intelligent age-friendly product characteristics on users' CUI: The mediating role of user experience perception. *E-Commerce Letters*. 2024;13:1353 [in Chinese].
- [51] Huang YC. Research on age-friendly design of mobile medical applications based on user experience element model. *Beauty and Times (Shang)*. 2024;(12):12–15 [in Chinese]. <https://doi.org/10.16129/j.cnki.mysds.2024.12.005>
- [52] Domínguez-Quintero AM, González-Rodríguez MR, Roldán JL. The role of authenticity, experience quality, emotions, and satisfaction in a cultural heritage destination. In: *Authenticity and Authentication of Heritage*. Routledge; 2021. p. 103–117.
- [53] Shen H, Chen QY, Zhang J. Research on the influence of virtual digital human anthropomorphism on tourists' QoE. *Journal of Central China Normal University*. 2025;59(3):358 [in Chinese].
- [54] Deng L, Turner DE, Gehling R, Prince B. User experience, satisfaction, and continual usage intention of IT. *European Journal of Information Systems*. 2010;19(1):60–75.
- [55] Sharabati AAA, Al-Haddad S, Al-Khasawneh M, Nababteh N, Mohammad M, Abu Ghoush Q. The impact of TikTok user satisfaction on continuous intention to use the application. *Journal of Open Innovation: Technology, Market, and Complexity*. 2022;8(3):125.
- [56] Chew HSJ, Achananuparp P. Perceptions and needs of artificial intelligence in health care to increase adoption: Scoping review. *Journal of Medical Internet Research*. 2022;24(1):e32939.
- [57] Khamaj A. AI-enhanced chatbot for improving healthcare usability and accessibility for older adults. *Alexandria Engineering Journal*. 2025;116:202–213.
- [58] Sun XX, Sun YN, Zhao YX, et al. Research on the influence of voice characteristics of AI medical voice assistants on elderly people's perceived credibility and usage intention. *Data Analysis and Knowledge Discovery*. 2025;9(7):104–117 [in Chinese].
- [59] Zhang X, Xue C. Exploring the role imagery preference of medical virtual agents developed through prompt engineering for the Chinese elderly: A mixed-methods study. In: *HCI 2025, LNCS 15819*; 2025. p. 267–285.
- [60] Singh K, Prabhu A, Kaur N. The impact and role of artificial intelligence (AI) in healthcare: Systematic review. *Current Topics in Medicinal Chemistry*. 2025;25(22):2635–2654.
- [61] Bergamin E, Roeser S. Technological mediation of emotional practices: The case of AI in healthcare. *Journal of Human-Technology Relations*. 2025;3:1–18.
- [62] Nicolas S, Agnieszka W. The personality of anthropomorphism: How the need for cognition and the need for closure define attitudes and anthropomorphic attributions toward robots. *Computers in Human Behavior*. 2021;122:106841.
- [63] Shi Z. Empowering Healthcare: Design-Driven AI Innovation and User Experience Optimization. *Big. D*. 2025;2(1):25–32. <https://doi.org/10.64504/big.d.v2i1.152>
- [64] Vyas D, van der Veer GC. APEC: A framework for designing experience. In: *Spaces, Places & Experience in HCI*; 2005. p. 1–4.
- [65] Lavie T, Tractinsky N. Assessing dimensions of perceived visual aesthetics of web sites. *International Journal of Human-Computer Studies*. 2004;60(3):269–298.
- [66] Khan ZF, Alotaibi SR. Applications of artificial intelligence and big data analytics in m-health: A healthcare system perspective. *Journal of Healthcare Engineering*. 2020;2020(1):8894694.
- [67] Morville P. A brief history of information. In: *Information Architecture: Designing Information Environments for Purpose*. 2004. p. 1.
- [68] Ahmed S, Aziz NA. Impact of AI on customer experience in video streaming services: A focus on personalization and trust. *International Journal of Human-Computer Interaction*. 2025;41(12):7726–7745.
- [69] Wattoo MU, Du J, Shahzad F, Kousar S. Shaping e-commerce experiences: Unraveling the impact of service quality on youth customer behavior in a developing nation. *Sage Open*. 2025;15(1):21582440241311786.
- [70] Reiter U, Brunnström K, De Moor K, Larabi MC, Pereira M, Pinheiro A, et al. Factors influencing quality of experience. In: *Quality of Experience: Advanced Concepts, Applications and Methods*. Cham: Springer International Publishing; 2014. p. 55–72.
- [71] Askari M, Klaver NS, van Gestel TJ, van de Klundert J. Intention to use medical apps among older adults in the Netherlands: Cross-sectional study. *Journal of Medical Internet Research*. 2020;22(9):e18080.
- [72] Griffiths JR, Johnson F, Hartley RJ. User satisfaction as a measure of system performance. *Journal of Librarianship and Information Science*. 2007;39(3):142–152.
- [73] Gao J, Opote AP, Jawad C, Zhan M. The influence of artificial intelligence chatbot problem solving on customers' continued usage intention in e-commerce platforms: An expectation-confirmation model approach. *Journal of Business Research*. 2025;200:115661.
- [74] Sun P, Li L, Hossain MS, Zabin S. Investigating students' behavioral intention to use ChatGPT for educational purposes. *Sustainable Futures*. 2025;9:100531.
- [75] Larivière B, Verleye K, De Keyser A, Koerten K, Schmidt AL. The service robot customer experience (SR-CX): A matter of AI intelligences and customer service goals. *Journal of Service Research*. 2025;28(1):35–56.
- [76] Babar Z, Paul R, Rahman MA, Barua T. A systematic review of human-AI collaboration in IT support services: Enhancing user experience and workflow automation. *Journal of Sustainable Development and Policy*. 2025;1(01):65–89.
- [77] Oksenberg L, Kalton G. New strategies for pretesting survey questions. *Journal of Official Statistics*. 1991;7(3):349.
- [78] Jackson DL. Revisiting sample size and number of parameter estimates: Some support for the N:q hypothesis. *Structural Equation Modeling*. 2003;10(1):128–141.
- [79] World Wide Web Consortium. Web Content Accessibility Guidelines (WCAG) 2.2. W3C; 2023. Available at: <https://www.w3.org/TR/WCAG22/> (accessed 2025).
- [80] International Organization for Standardization. ISO 9241-210:2019. Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems. Geneva: ISO; 2019.
- [81] Nielsen J. *Usability Engineering*. San Francisco: Morgan Kaufmann; 1993.
- [82] International Telecommunication Union. ITU-T Recommendation E.800: Definitions of terms related to quality of service. Geneva: ITU; 2008.
- [83] National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework (AI RMF 1.0). Gaithersburg: NIST; 2023.

- [84] Picard RW. *Affective Computing*. Cambridge: MIT Press; 1997.
- [85] Brislin RW. Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*. 1970;1(3):185–216.
- [86] Doshi-Velez F, Kim B. Towards a rigorous science of interpretable machine learning. *arXiv:1702.08608*; 2017.
- [87] Polit DF, Beck CT. The content validity index: Are you sure you know what's being reported? A critique and recommendations. *Research in Nursing & Health*. 2006;29(5):489–497.
- [88] Hayes AF. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. 2nd ed. New York: Guilford Press; 2018.
- [89] World Health Organization. *Integrated care for older people: Guidelines on community-level interventions to manage declines in intrinsic capacity*. Geneva: WHO; 2017.
- [90] Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999;6(1):1–55.