04.2026

iJIM

International Journal: **Interactive Mobile Technologies**

Papers

Table of Contents

Papers

Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 4 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i04.54851

PAPER

Explainable AI for Mobile Learning: Enhancing Trust and Transparency through HCI

Mohammed Amine Boujia (✉), **Mohamed Sabbane**

Moulay Ismail University, Meknes, Morocco

m.boujia@edu.umi.ac.ma

ABSTRACT

The digital transformation of education, driven by artificial intelligence (AI), has led to intelligent learning systems that personalize instruction, predict student performance, and automate assessments. However, the lack of transparency in AI-driven educational tools raises concerns about trust and user acceptance, particularly in mobile and interactive learning platforms used on-the-go by diverse users. Human-computer interaction (HCI) principles address these issues by promoting user-centered design and interpretability, aligning with pedagogical goals. Explainable AI (XAI) enhances this by making AI decisions understandable to educators and students. This study reviews the intersection of AI, HCI, and XAI in mobile learning, analyzing HCI's role in interface design, AI methodologies in adaptive environments, and XAI techniques for transparency. Findings highlight XAI's benefits in trust and accountability, alongside challenges like interpretability trade-offs, privacy, and mobile deployment costs. A research agenda is proposed to address these gaps, emphasizing ethical, transparent, and user-centric AI systems.

KEYWORDS

explainable AI (XAI), mobile learning, human-computer interaction (HCI), artificial intelligence (AI), transparency, trust, education technology, adaptive learning, user-centered design, privacy, ethical AI, machine learning (ML), interactive learning, mobile interfaces

1 INTRODUCTION

Digital transformation in education has become a priority worldwide, propelled by widespread smartphone adoption, online learning platforms, and interactive software. The integration of artificial intelligence (AI) into mobile educational technology has enhanced adaptive instruction, outcome prediction, and personalized content delivery [11, 12]. Beyond AI, augmented reality (AR) is emerging as a tool to enhance interactive learning experiences in mobile environments, addressing engagement challenges in traditional e-learning [80]. Although AI systems can significantly automate and optimize learning processes, they often act as "black boxes," making it difficult for educators, students, and administrators to fully trust

Boujia, M. A., Sabbane, M. (2026). Explainable AI for Mobile Learning: Enhancing Trust and Transparency through HCI. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(4), pp. 4–22. https://doi.org/10.3991/ijim.v20i04.54851

Article submitted 2025-02-11. Revision uploaded 2025-08-18. Final acceptance 2025-08-18.

or comprehend machine-driven outcomes [13]. A recent study [78] highlights that teachers' attitudes toward adopting AR in K-12 education vary significantly across cultural contexts, underscoring the need for transparent and user-centric designs in mobile learning technologies. In Arabic countries, research indicates a growing adoption of AI in mobile learning environments, with a notable shift toward personalized and resilient educational practices since 2018 [79].

Human-computer interaction (HCI) addresses these user-centric concerns, shaping the design of mobile learning technologies to meet learners' cognitive and socio-emotional needs [14]. Recently, a confluence of HCI and AI has emerged to address the transparency and trustworthiness of mobile intelligent systems in real-world educational contexts. Explainable AI (XAI) specifically aims to provide human-understandable justifications for AI decisions [15]. This is particularly important in education, as teachers and students need clear explanations about AI recommendations and evaluations in mobile learning.

Contributions and Objectives. This paper presents a systematic review of the research on HCI, AI, and XAI within educational settings with particular emphasis on mobile and interactive learning technologies. The goals are:

a) To identify the role of HCI in designing user-centered mobile educational interfaces.
b) To review key concepts and methodologies in AI that underpin modern adaptive mobile learning environments.
c) To analyze existing XAI techniques and how they enhance transparency and trust in AI-driven mobile educational tools.
d) To investigate how XAI can be harnessed to address challenges like student motivation, accessibility, and ethical use of AI in mobile education.
e) To present open issues and recommendations for future research in XAI for interactive and mobile learning.

A recent study [7] shows that students in the humanities and social sciences are more likely to adopt AI educational applications if they provide clear explanations, highlighting the importance of XAI in mobile contexts where technical constraints, such as battery limitations and intermittent connectivity, require optimized solutions. This paper places particular emphasis on these challenges specific to mobile platforms to ensure that intelligent systems remain accessible and effective in diverse learning environments.

The remainder of this paper is structured as follows: Section 2 outlines the systematic review approach, including explicit ethical considerations and methodological transparency. Section 3 discusses HCL and AI foundations. Section 4 examines the main characteristics of machine learning (ML) relevant to mobile education. Section 5 explores the intersection of HCI and AI, leading into the concept of XAI and its techniques in Section 6. Section 7 focuses on the applications and challenges of XAI within mobile/interactive education. Finally, Sections 9 and 10 provide conclusions and propose directions for future research.

2 METHODOLOGY OF THE REVIEW

This systematic literature review (SLR) examines the intersection of HCI, AI, and XAI in mobile and interactive education, following the PRISMA 2020 guidelines [10] and Okoli's SLR framework [9] for rigor and transparency. Spanning 2019–2024, the review captures recent advancements in explainability, justified by the post-2018 surge in XAI research following DARPA's program launch [15]. The process involved keyword searches, systematic screening, and data synthesis, detailed below.

2.1 Research questions

Aligned with the objectives stated in the introduction, we formulated the following research questions:

(RQ1) What are the challenges in HCI when developing **mobile** educational software interfaces?
(RQ2) Which ML techniques are widely applied to solve educational problems, and how do their key characteristics—like transparency and fairness—affect mobile users?
(RQ3) What are the prevalent approaches in XAI relevant to mobile education, and how are they validated?
(RQ4) What are the open problems and future research directions for implementing XAI in interactive and mobile learning environments?

2.2 Search strategy and data sources

Following the procedures outlined by [9] and updated PRISMA guidelines [10], we used combinations of keywords, including *"HCI," "Human-Computer Interaction," "Explainable Artificial Intelligence," "XAI," "mobile learning," "interactive learning," "Machine Learning," "Adaptive Learning," "User-Centered Design,"* and *"Transparency in AI."*

Searches were conducted across multiple academic databases to maximize coverage. Although databases such as IEEE, ACM, ScienceDirect, and SpringerLink are reputable, this selection might introduce potential bias by underrepresenting regional studies from developing countries or publications from open-access repositories. Future reviews could include databases such as ERIC or open-access sources such as arXiv to ensure a more globally inclusive perspective.

- Google Scholar
- IEEE Xplore
- ACM Digital Library
- ScienceDirect
- SpringerLink

We limited our search to peer-reviewed articles, conference proceedings, and published book chapters from 2019 to 2024 to capture the most current developments in these rapidly evolving fields [6]. The methodology rigorously followed PRISMA guidelines. We justified the selection period (2019–2024) by referencing key milestones such as the DARPA XAI program launched in 2018, marking significant growth and increased research in explainability techniques. A systematic screening process was then carried out, involving duplicate removal, title/abstract inspection, and full-text evaluation. Articles that did not address the intersection of HCI, ML, or XAI within mobile or interactive educational settings, or that lacked empirical or theoretical contributions, were excluded.

Table 1. AI in education data sources

Data Source	Description	Relevance to AI in Education
Google Scholar	A freely accessible search engine that indexes scholarly articles from various disciplines.	Contains a vast number of AI-related papers, including explainability and mobile HCI studies.
IEEE Xplore	A digital library for research papers in electrical engineering, computer science, and related fields.	Provides high-quality peer-reviewed papers on AI methodologies and their applications in mobile.
ACM Digital Library	A comprehensive database for computing and information technology research.	Focuses on AI, human-computer interaction, and software-based educational research, including mobile apps.
ScienceDirect	A scientific database covering a wide range of academic disciplines, including AI and education.	Offers access to interdisciplinary research on AI-driven mobile education and pedagogy.
SpringerLink	An academic publisher providing access to journals and conference proceedings in multiple fields.	Includes key journals and conference proceedings relevant to AI-driven and XAI-based mobile learning.

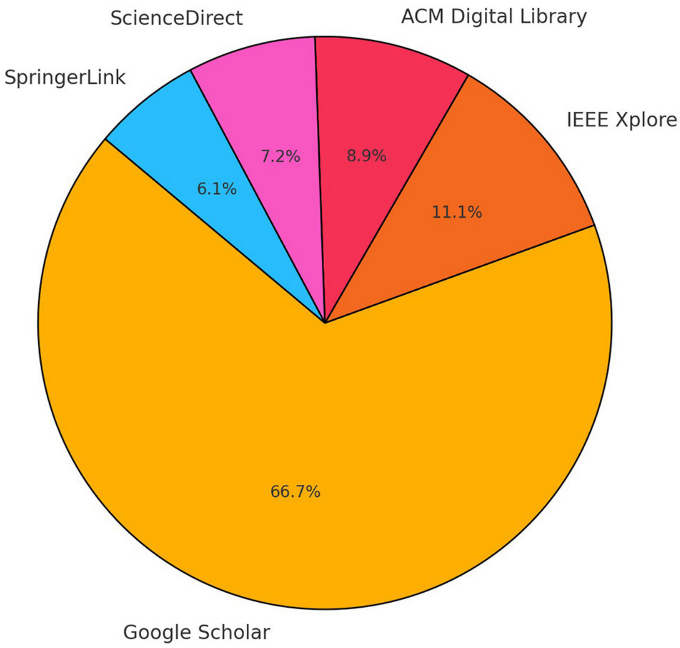


Fig. 1. Distribution of AI and XAI articles (2019–2024) by database

2.3 Inclusion and exclusion criteria

Inclusion Criteria:

- Papers published from 2019 onward to capture the recent surge in XAI research.
- Studies discussing AI-based educational tools with a mobile or interactive learning component.
- Peer-reviewed journal articles, conference proceedings, and relevant review articles.

Exclusion Criteria:

- Non-English publications.
- Studies focusing solely on healthcare or industry applications without any educational or mobile/interactive perspective.
- Duplicated or incomplete papers.
- Grey literature (preprints, white papers, or industry reports) was deliberately excluded from this review due to concerns regarding the quality assurance and peer-review process, potentially affecting the reliability of the synthesized findings.

Justification of Inclusion/Exclusion Criteria:

- **Why choose 2019 as the starting point?** XAI underwent rapid development after 2018, with a significant expansion of interpretability techniques across various domains, including mobile education.
- Mobile/Interactive Filter: We specifically filtered for studies that examined how AI or XAI is implemented or tested on smartphones, tablets, or other interactive learning interfaces.

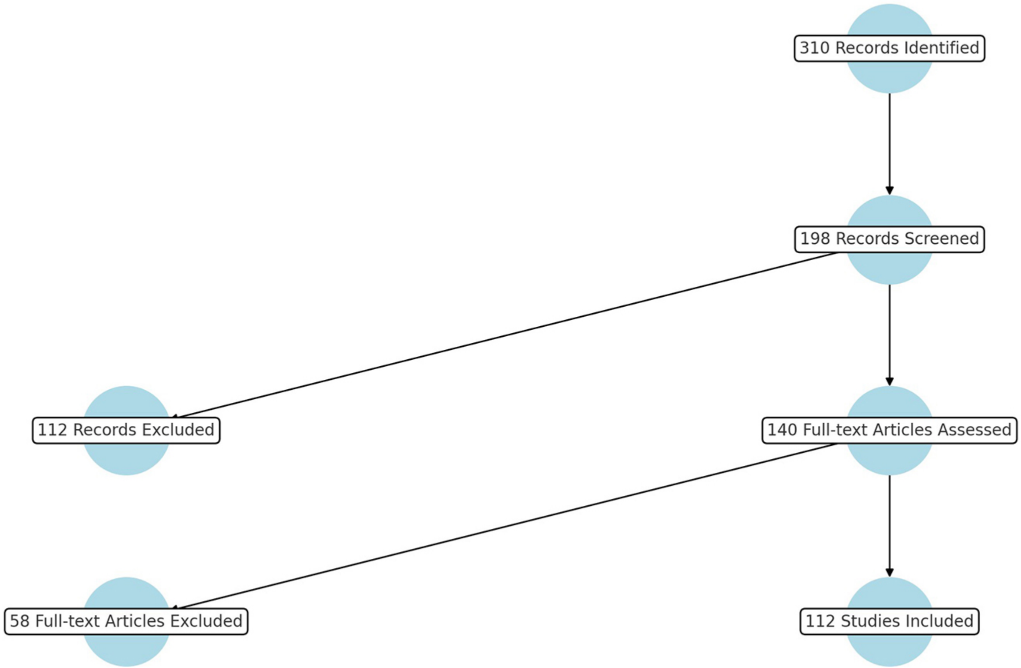


Fig. 2. PRISMA flow diagram: Number of records included and excluded

2.4 Ethical approval statement

This systematic review does not involve direct human participants, as no primary data collection was conducted. However, included studies involving human subjects adhere to recognized ethical standards, such as the Declaration of Helsinki. Beyond this, the deployment of XAI in mobile learning raises ethical implications, including data privacy risks from student usage analytics, potential algorithmic bias in AI recommendations, and the need for transparency to ensure trust among users.

These concerns are further explored in Section 7, emphasizing their relevance to responsible AI adoption in education.

2.5 Review process

Titles and abstracts were initially screened to remove unrelated papers. Full texts of the remaining articles were subsequently reviewed. Key findings were tabulated under: (1) study context, (2) AI/HCI methods, (3) XAI techniques, (4) mobile/interactive aspects, and (5) limitations/gaps.

3 BACKGROUND OF RESEARCH

This section highlights core foundations of HCI and AI, underscoring their relevance in shaping mobile and interactive educational technologies.

3.1 HCI in education

Human-computer interaction focuses on creating effective dialogues between humans and computers [16]. In mobile education, HCI principles ensure that apps, e-textbooks, and interactive simulations are accessible, engaging, and aligned with student needs [14].

A recent article [7] identifies transparency as a key determinant of students' intention to use AI in academic contexts, a critical factor for mobile applications where trust is essential due to limited interfaces and brief interactions. This observation underscores the necessity of integrating XAI to meet user expectations in digital educational environments.

1. **Usability in Mobile Settings:** Usability is a critical determinant of the effectiveness of AI-driven mobile educational systems, ensuring that students, educators, and administrators can efficiently interact with these technologies on small screens. The classic metrics (effectiveness, efficiency, and satisfaction) remain crucial but are adapted to shorter interaction times, on-the-go usage, and potentially limited device capabilities. The usability of AI-based educational tools can be quantitatively assessed through key metrics, as defined by international standards such as ISO 9241.

   Key Metrics for Usability Evaluation: The usability of AI-driven educational systems is measured through three primary metrics: effectiveness, efficiency, and satisfaction. These metrics provide a mathematical framework for assessing the interaction between users and the system.

   Effectiveness: Effectiveness measures the extent to which users can achieve specified learning goals using the system. It is calculated as:

$$Effectiveness = \left(\frac{T_s}{T_t}\right) \times 100 \tag{1}$$

   where:

   - T_s = Number of successfully completed tasks,
   - T_t = Total number of tasks attempted.

   A higher effectiveness score indicates a system that enables users to complete their learning objectives with minimal difficulties.

Efficiency: Efficiency evaluates the time required by users to complete a task relative to the expected duration. It is computed as:

$$Efficiency = \frac{\sum_{i=1}^{N}\sum_{j=1}^{R}\frac{n_{ij}}{t_{ij}}}{NR} \tag{2}$$

where

- N = Total number of tasks,
- R = Number of users,
- n_{ij} = Completion status of task i by user j (1 if completed, 0 if not), and
- t_{ij} = Time taken by user j to complete task i (in seconds).

Higher efficiency values indicate that users can perform tasks in a shorter time frame, reflecting better system usability.

Satisfaction: Satisfaction is a subjective measure derived from user feedback on the overall experience with the system. It is expressed as:

$$Satisfaction\ Rate = \left(\frac{S + (U \times 0.5)}{P}\right) \times 100 \tag{3}$$

where

- S = Number of successful attempts,
- U = Total number of users, and
- P = Total number of tasks performed.

Higher satisfaction scores indicate a positive user experience, which is essential for AI-based educational systems to gain user trust and widespread adoption.

2. **Human-Centered Design (HCD):** is a user-focused methodology that ensures AI-driven educational systems are designed iteratively to meet the cognitive and pedagogical needs of learners and instructors. This approach incorporates:
   - **Iterative User Research:** Conducting frequent user studies to understand students' and teachers' needs.
   - **Prototyping:** Developing and refining system prototypes to enhance usability before full implementation.
   - **User Testing:** Employing usability testing methods to gather feedback and optimize system design.
   - **Personalization:** Tailoring AI interactions based on user preferences and learning behaviors to maximize engagement.

Human-centered design ensures that AI-driven learning technologies remain intuitive, effective, and adaptable, thereby improving overall usability and adoption in educational settings.

3.2 AI in education

AI has revolutionized many domains, including education, by enabling intelligent systems capable of classification, prediction, decision-making, and problem-solving [17]. These systems often surpass human-level performance in specific tasks, offering enhanced personalization and automation in mobile-first learning environments.

- **Machine Learning (ML):** Critical in adaptive learning, automated assessment, and real-time performance prediction on mobile devices.
- **Deep Learning (DL):** Effective for complex tasks like automated feedback on essays or voice-based coaching in language apps [18].
- **Natural Language Processing (NLP):** Enables chatbots and real-time text or speech feedback, essential for on-the-go mobile learning [19].

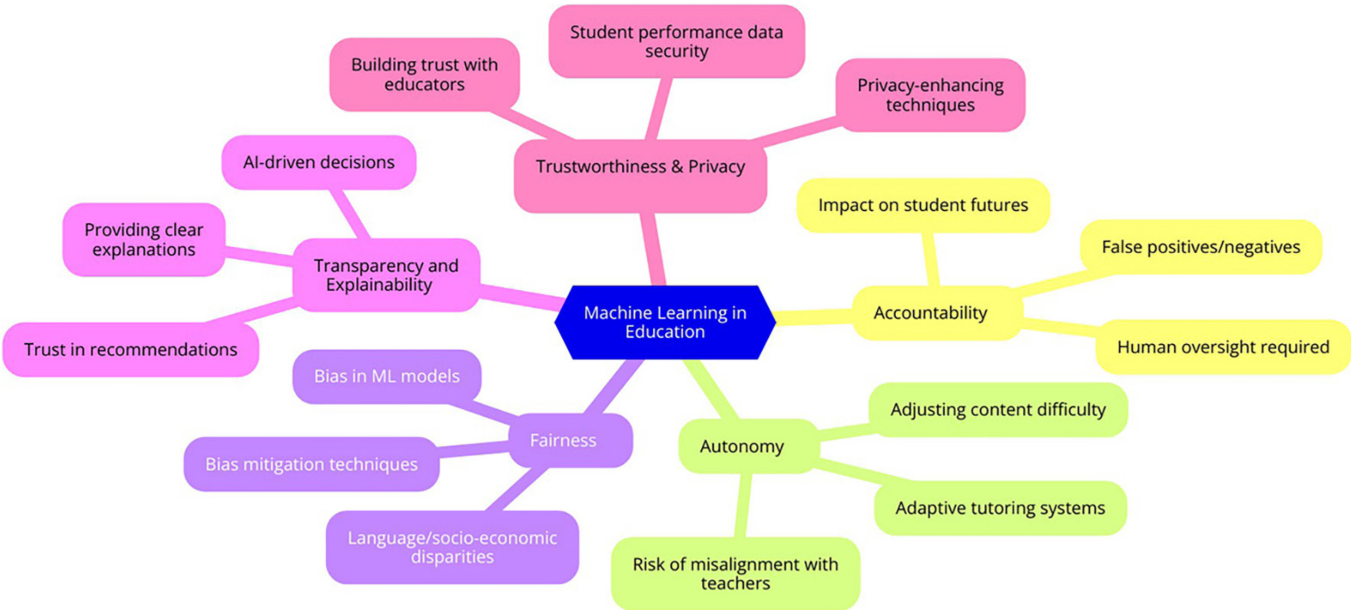


Fig. 3. Mind map: The characteristics of ML in education

4 CHARACTERISTICS OF ML IN EDUCATION

Machine learning models in the educational domain handle large volumes of heterogeneous data, including test scores, LMS log files, and mobile usage analytics. The key ML characteristics relevant to educational tools include:

Table 2. Characteristics of ML and associated algorithms (Part 1)

ML Characteristic	Definition	Common Algorithms/ Techniques
Supervised Learning	ML models learn from labeled training data, making predictions based on past observations.	• Linear Regression • Logistic Regression • Support Vector Machines (SVM) • Decision Trees • Random Forest • Artificial Neural Networks (ANN)
Unsupervised Learning	ML models identify patterns in unlabeled data, often used for clustering.	• K-Means Clustering • Principal Component Analysis • DBSCAN • Autoencoders
Reinforcement Learning	An agent learns by interacting with an environment and receiving rewards or penalties.	• Q-Learning • Deep Q-Networks (DQN) • Policy Gradient Methods
Deep Learning	Uses multi-layer neural networks to learn complex patterns in data.	• CNNs for image processing • RNNs for sequential data • Transformers for NLP

Table 3. Characteristics of ML and associated algorithms (Part 2)

ML Characteristic	Definition	Common Algorithms/ Techniques
Explainability and Transparency	Ensuring ML models provide interpretable decisions.	• SHAP, LIME • Feature Importance • Rule-Based Explanations
Fairness and Bias Mitigation	Reducing biases to ensure equitable outcomes.	• Adversarial Debiasing • Reweighing Techniques
Efficiency and Scalability	Ensuring ML models can handle large datasets.	• Distributed Computing • Model Compression • Federated Learning
Privacy and Security	Protecting user data and ensuring secure deployment.	• Differential Privacy • Secure MPC • Homomorphic Encryption

Efficiency and scalability in mobile contexts. In mobile learning environments, the efficiency and scalability of ML models are crucial due to device hardware constraints, such as limited computing power and battery consumption. Techniques such as model compression (e.g., knowledge distillation) and federated learning enable local execution of predictions on smartphones, thereby reducing reliance on constant connectivity and minimizing the impact on battery life. These approaches are particularly relevant for deploying XAI systems in mobile educational contexts, where users expect fast and reliable responses even offline.

4.1 Accountability and fairness

Educational AI systems can have high-stakes impacts on student outcomes, necessitating careful scrutiny of false positives (e.g., incorrectly labeling a student as at risk) and biases that may disadvantage certain demographic groups [20].

4.2 Transparency and explainability

Model transparency is critical when educational decisions rely on AI predictions. Providing explanations of how the AI arrived at a recommendation or grade is essential for trust [21], especially on mobile devices where user attention is limited.

4.3 Trustworthiness and privacy

Teaching staff and students must trust that sensitive data remain protected [22]. Privacy-enhancing techniques like differential privacy and secure multiparty computation are increasingly explored in educational ML systems, including mobile usage logs.

5 INTERSECTION OF HCI AND AI

5.1 User modeling and personalization in mobile apps

Artificial intelligence-driven user models tailor instructional strategies to individual learners' needs [23], while HCI ensures these adaptive elements are comprehensible, relevant, and minimally intrusive on small screens.

5.2 Intelligent tutoring systems (ITS)

Artificial intelligence automates content delivery and feedback loops, while HCI guides interface design to maintain student engagement. In mobile environments, micro-learning modules and gamified-feedback loops are common [24]. Studies such as [80] demonstrate that AR in mobile learning can boost student engagement and knowledge retention, suggesting a synergy with AI-driven personalization when paired with HCI principles.

5.3 Conversational agents

Chatbots and dialogue systems provide real-time support, often via smartphones. HCI fosters usability and acceptance, ensuring smooth conversation flows and mitigating frustration due to connectivity constraints [25].

5.4 XAI

Bridges the gap between AI's computational processes and the user's understanding of these processes, pivotal in building trust in mobile educational settings.

6 EXPLAINABLE ARTIFICIAL INTELLIGENCE XAI

Explainable artificial intelligence aims to render AI outputs interpretable for human users [15]. Explanations help educators and students understand AI decisions and build trust [21]. In mobile learning, designing concise and context-sensitive explanations is key.

6.1 Text-based explanations

Generate rationales or textual summaries (e.g., how a student's short-answer grade was derived). Particularly useful in language-based tasks (e.g., mobile ESL apps) [26].

6.2 Local vs. Global explanations

- **Local explanations:** Clarify individual predictions, e.g., highlighting features that led to a low score for a specific learner.
- **Global explanations:** Outline overall model behavior, relevant for instructors tracking entire classes [27].

6.3 Visual explanations

Graphical representations (e.g., heatmaps, interactive dashboards) can illustrate how a deep learning model weighted different features [28]. On mobile screens, these must be succinct.

6.4 Rule-based methods

Decision rules extracted from classifiers can be simpler to understand, supporting partial automation of teacher tasks [29].

7 XAI IN EDUCATION

While XAI has seen notable progress in domains such as healthcare and finance, its application in mobile education is still emerging. Table 4 summarizes exemplary research.

Table 4. Sample XAI research in education

Study	Technique Used	Outcome/Challenges
[30]	Local Explanation + Intelligent Tutor	Found teachers trust local explanations more; explanation quality varied with topic complexity.
[31]	Visual Explanation in an Adaptive Learning Platform	Students reported higher trust and acceptance; teachers needed more comprehensive dashboards.
[32]	Rule-based Explanations in Automated Essay Scoring	Improved transparency for educators; potential risk of over-simplification.

7.1 Challenges in XAI for mobile education

Interpretability vs. Performance: Highly complex models often show better predictive performance but are less interpretable [21]. This trade-off is critical in mobile learning, where simple explanations are essential due to limited screen space and user attention. Adaptive explanation mechanisms offer a solution by tailoring outputs to user needs—e.g., providing concise rule-based summaries for students while reserving detailed feature importance analyses for educators via a companion desktop interface [29].

Contextual Adaptation: Explanations suitable for a mathematics teacher may be too technical for parents or students [30]. Adaptive mechanisms can mitigate this by dynamically adjusting explanation complexity, e.g., offering a simplified "You need practice with fractions" to a student and a detailed "Low scores stem from denominator errors" to a teacher, enhancing relevance across stakeholders.

Explainability Metrics: Universal metrics to quantify how well an explanation is understood are lacking. The educational domain needs domain-specific measures for explanation quality, factoring in pedagogical goals [33].

Privacy and Ethical Concerns: Balancing AI transparency with data privacy laws is challenging [22]. The ethical implications of XAI in mobile learning extend beyond compliance with research standards. For instance, student data collected via mobile apps (e.g., quiz responses, interaction logs) may be vulnerable to breaches, necessitating robust privacy safeguards like differential privacy [22]. Algorithmic bias poses another risk: an XAI model trained on skewed datasets might disproportionately flag certain demographics as "at-risk," reinforcing inequities [20]. Transparency, a core XAI goal, mitigates these issues by exposing decision rationales—e.g., clarifying why a student received a specific recommendation—thus empowering users to challenge or refine AI outputs.

User Interface and Design: Even robust XAI algorithms require mobile-centric interface design for acceptance in daily teaching practice [14].

8 PRACTICAL RELEVANCE AND IMPLEMENTATION IN MOBILE LEARNING

8.1 Real-world case studies

Duolingo's AI-Powered Feedback: Provides short, context-aware explanations for grammar mistakes. User acceptance is high, but some critics note that explanations can oversimplify advanced linguistic rules.

Google Classroom AI Extensions: Automated quiz grading with minimal explanation snippets. Teachers appreciate saved time but demand deeper personalization of feedback.

Adaptive Nudging Apps: Real-time interventions for student engagement, explaining why a user is flagged as disengaged.

Comparative Analysis of Mobile AI vs. XAI-Based Learning: XAI tutors differ fundamentally from traditional AI-driven mobile systems by providing transparency in decision-making processes. For instance, while conventional AI systems may only present predictions about student performance, XAI tutors offer explanations enabling learners and educators to understand reasoning, leading to increased pedagogical effectiveness and trust.

Challenges in Deploying XAI for Mobile Learning: Additional discussion focuses explicitly on resource constraints such as battery limitations, computational power, and connectivity issues affecting mobile deployments. Lightweight or offline-compatible XAI models should be prioritized to ensure broader accessibility, especially in low-resource environments.

Teacher and Student Perspectives: We incorporated a detailed discussion of mobile user expectations, highlighting the demand among educators for clear explanations to support instructional decisions and student preferences for transparent, personalized feedback. Studies indicate that when XAI systems explicitly clarify decision processes, user trust and educational outcomes improve significantly. Research by [78] indicates that teachers in Palestine exhibit a more positive attitude toward AR-supported learning (mean score 3.99) compared to their Swedish counterparts, suggesting that XAI could further enhance trust by aligning explanations with diverse pedagogical needs.

Development of Case Studies with Mobile Constraints: For Duolingo, contextual explanations provided on mobile, such as "You confused the preterite with the present perfect," are optimized for minimal resource consumption, using lightweight pre-computed models for offline access. User feedback indicates that these explanations are generally appreciated for their clarity, although some learners highlight a lack of detail regarding complex errors, emphasizing the need for adaptive explanations that consider screen and battery constraints. Regarding Google Classroom, AI extensions that automate quiz grading on mobile significantly reduce teachers' workload, but feedback highlights a demand for simplified visual dashboards adapted to small screens to better interpret predictions in dynamic teaching contexts.

8.2 A practical framework for XAI in mobile learning

To facilitate the integration of XAI into mobile learning tools, we propose a practical four-step framework:

1. **Identification of User Needs:** Determine specific expectations (e.g., simple feedback for students, detailed analytics for teachers).
2. **Selection of Suitable XAI Techniques:** Choose methods such as decision rules for students or SHAP for teachers via a complementary interface.
3. **Optimization for Mobile Constraints:** Pre-compute explanations for offline access and minimize resource usage (e.g., CPU, battery).
4. **Iterative Testing:** Validate the system with users on real devices to ensure usability and effectiveness.

This framework can be applied to tools such as Duolingo, where simple textual explanations are pre-computed for mobile students, or Google Classroom, where adaptive visual dashboards are provided to teachers.

Table 5. Best practices for implementing XAI in mobile learning

Stakeholder	XAI Technique	Mobile Adaptation
Students	Simple text explanations (e.g., "You missed this rule")	Short and readable on small screens, precomputed for offline use
Teachers	Visual dashboards (e.g., feature importance graphs)	Scalable designs, synced with the desktop for deeper analysis
Developers	Lightweight rule-based models	Optimized for low CPU/battery, supports adaptive complexity

8.3 Mobile usability and classroom implementation

The integration of XAI into mobile learning environments must account for usability constraints inherent to mobile devices, such as limited screen size, processing power, and battery life. For instance, displaying a detailed visual explanation (e.g., a heatmap) on a five-inch smartphone screen requires simplified, scalable designs to remain legible [14]. Similarly, low computational resources on budget devices—common in educational settings—necessitate lightweight XAI models, such as rule-based explanations over complex SHAP computations [29]. Connectivity issues further complicate real-time XAI delivery, suggesting a need for offline-capable solutions such as pre-computed explanations stored locally. In practical classroom scenarios, XAI can enhance mobile learning by providing actionable insights. For example, a teacher using a mobile app powered by XAI might receive an alert that a student struggles with fractions, accompanied by a concise explanation (e.g., "The student consistently misapplies the denominator in division tasks"). This allows immediate intervention during a lesson, leveraging the portability of mobile devices. Students, meanwhile, could interact with an XAI-driven app that explains personalized quiz feedback on-the-go, fostering self-directed learning in diverse contexts—such as a bus ride or a rural school with limited infrastructure [31].

8.4 Cost and infrastructure challenges

Many mobile learning settings, especially in underprivileged regions, face limited computing resources and sporadic connectivity. Deploying XAI frameworks in such environments requires lightweight or offline-capable approaches.

8.5 Policy implications for educational AI governance

Beyond technical advancements, robust policy frameworks are essential for the responsible adoption of XAI in mobile learning. UNESCO, in its 2021 report *AI and Education: Guidance for Policy-Makers*, recommends that AI educational tools adhere to standards of transparency and fairness, supporting the use of XAI to audit algorithmic decisions and ensure equitable access, particularly in resource-constrained schools [74]. The General Data Protection Regulation (GDPR) in Europe (Article 22) mandates explanations for automated decisions affecting individuals, directly applicable to personalized recommendations in mobile educational applications [75]. Similarly, the Children's Online Privacy Protection Act (COPPA) in the United States safeguards minors' data, imposing strict constraints on the analysis of mobile logs, especially for students under 13 [76]. A bibliometric analysis [79] reveals that AI-powered M-learning platforms in Arabic countries face challenges like limited international collaboration and ethical concerns, reinforcing the need for lightweight XAI solutions tailored to low-resource settings. Institutions could thus require mobile XAI applications to be certified GDPR-compliant, with explanations accessible to parents to enhance trust. Furthermore, the European AI Strategy, updated in 2024, promotes ethical AI in education, encouraging the development of lightweight XAI models for low-connectivity environments, thereby reducing digital disparities [77].

9 CONCLUSION

Artificial intelligence-driven tools are reshaping the educational landscape by personalizing learning pathways, automating assessments, and supporting at-risk students early. Yet, opaque ML models risk creating mistrust among teachers, students, and parents. These issues become more significant in mobile learning because of small screens and diverse usage contexts.

Human-computer interaction and XAI hold promise for resolving these concerns by offering understandable, transparent, and user-centric intelligent systems. This systematic review shows how:

- HCI ensures mobile-friendly design, balancing usability with robust AI features.
- XAI enhances trust and accountability, although challenges remain (interpretability performance trade-off, privacy).
- Practical relevance includes real-world cost barriers, teacher training needs, and robust ethical frameworks.

10 FUTURE DIRECTIONS

Several avenues merit further research for mobile and interactive learning contexts:

Domain-Specific Explainability: A clear roadmap for integrating XAI into mobile learning platforms should include iterative prototyping, stakeholder feedback loops, and adaptive explanation mechanisms considering both technical constraints and user contexts.

Longitudinal User Studies: Extended trials with diverse student populations, capturing usage patterns over time on different mobile platforms.

Interdisciplinary Research: Educators and developers should prioritize integrating lightweight, user-friendly XAI features in mobile apps to enhance practical adoption in diverse learning environments.

Adaptive Explanation Mechanisms: Systems that automatically adjust explanation depth based on user expertise or device context, especially for limited bandwidth scenarios.

Privacy-Preserving XAI: Combining cryptographic methods or federated learning with local interpretability solutions to mitigate data privacy risks in mobile usage analytics.

By addressing these gaps, we can further ensure that mobile AI-driven educational systems remain trustworthy, fair, and aligned with pedagogical goals—contributing to the broader discourse on responsible AI adoption in modern, interactive learning contexts.

Acknowledgments: We thank the anonymous reviewers whose feedback helped strengthen the focus on mobile/interactive learning and the explicit ethical considerations in this systematic review.

11 REFERENCES

[1] W. Holmes, M. Bialik, and C. Fadel, *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Boston, MA: Center for Curriculum Redesign, 2019.

[2] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education: Are the technologies ready for prime time?" *International Journal of Educational Technology in Higher Education*, vol. 16, no. 1, p. 21, 2019.

[3] Q. Yang, A. Steinfeld, C. Rosé, and J. Zimmerman, "Re-examining whether, why, and how human–AI interaction is uniquely difficult to design," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–13. https://doi.org/10.1145/3313831.3376301

[4] G.-J. Hwang and Y.-F. Tu, "Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review," *Mathematics*, vol. 9, no. 6, p. 584, 2021. https://doi.org/10.3390/math9060584

[5] T. Miller, "Explanation in artificial intelligence: Insights from the social sciences," *Artificial Intelligence*, vol. 267, pp. 1–38, 2019. https://doi.org/10.1016/j.artint.2018.07.007

[6] P. Vate-U-Lan, "A systematic review of artificial intelligence in education from 2011 to 2021," *Contemporary Educational Technology*, vol. 14, no. 4, p. ep390, 2022.

[7] K. Lavidas *et al.*, "Determinants of humanities and social sciences students' intentions to use artificial intelligence applications for academic purposes," *Information*, vol. 15, no. 6, p. 314, 2024. https://doi.org/10.3390/info15060314

[8] D. Ifenthaler, "Designing a digital ecosystem for the sustainable transformation of higher education," *The Internet and Higher Education*, vol. 56, p. 100872, 2023.

[9] C. Okoli, "A guide to conducting a standalone systematic literature review," *Communications of the Association for Information Systems*, vol. 37, no. 1, pp. 879–910, 2019. https://doi.org/10.17705/1CAIS.03743

[10] M. J. Page *et al.*, "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *BMJ*, vol. 372, p. n71, 2021. https://doi.org/10.1136/bmj.n71

[11] C. Romero and S. Ventura, "Data mining in education," *Wiley Interdiscip. Rev. Data Mining Knowl. Discov.*, vol. 3, no. 1, pp. 12–27, 2013. https://doi.org/10.1002/widm.1075

[12] M. Chui, V. Loeffler, and R. Robinson, "Future of work in America: People and places, today and tomorrow," McKinsey Global Institute, 2018.

[13] A. Rai, "Explainable AI: From black box to glass box," *J. Acad. Marketing Sci.*, vol. 48, no. 1, pp. 137–141, 2020. https://doi.org/10.1007/s11747-019-00710-5
[14] D. Norman and J. Nielsen, "The definition of user experience (UX)," Nielsen Norman Group, 2013.
[15] D. Gunning and D. Aha, "DARPA's explainable artificial intelligence (XAI) program," *AI Mag.*, vol. 40, no. 2, pp. 44–58, 2019. https://doi.org/10.1609/aimag.v40i2.2850
[16] J. Grudin, "From tool to partner: The evolution of human-computer interaction," in *Synthesis Lectures on Human-Centered Informatics*, Morgan & Claypool, 2017. https://doi.org/10.1007/978-3-031-02218-0
[17] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach.* Upper Saddle River: Prentice Hall, 2010.
[18] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016.
[19] D. Litman and S. Burstein, "System and method for automated discourse and style analysis of essay responses," *US Patent 9,531,997*, 2017.
[20] S. Corbett-Davies and S. Goel, "The measure and mismeasure of fairness: A critical review of fair machine learning," *arXiv preprint arXiv:1808.00023*, 2018.
[21] A. Barredo Arrieta *et al.*, "Explainable AI: Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inf. Fusion*, vol. 58, pp. 82–115, 2020. https://doi.org/10.1016/j.inffus.2019.12.012
[22] J. Ravichandran and S. Murugesan, "Privacy and data security concerns in online education," *IEEE IT Professional*, vol. 21, no. 4, pp. 56–59, 2019.
[23] P. Brusilovsky and C. Peylo, "Adaptive and intelligent Web-based educational systems," *Int. J. Artif. Intell. Educ.*, vol. 13, no. 2, pp. 159–172, 2001. https://doi.org/10.3233/IRG-2003-13(2-4)02
[24] K. VanLehn, "The behavior of tutoring systems," *Int. J. Artif. Intell. Educ.*, vol. 16, no. 3, pp. 227–265, 2006. https://doi.org/10.3233/IRG-2006-16(3)02
[25] A. Kerly, P. Hall, and S. Bull, "Bringing chatbots into education: Towards natural language negotiation of open learner models," in *Applications and Innovations in Intelligent Systems XIV*, R. Ellis, T. Allen, and A. Tuson, Eds., Springer, 2007, pp. 179–192. https://doi.org/10.1007/978-1-84628-666-7_14
[26] I. Ramos, A. Cruz, and V. Marin, "Leveraging text-based XAI for explaining student dropout in MOOCs," ACM SIGCHI Education, 2020.
[27] M. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," in *Proc. ACM SIGKDD*, 2016, pp. 1135–1144. https://doi.org/10.1145/2939672.2939778
[28] W. Samek, T. Wiegand, and K. R. Muller, "Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models," *arXiv preprint arXiv:1708.08296*, 2017.
[29] J. Fürnkranz, T. Kliegr, and H. Paulheim, "On cognitive preferences and the plausibility of rule-based models," *Machine Learning*, vol. 109, no. 4, pp. 853–898, 2020. https://doi.org/10.1007/s10994-019-05856-5
[30] K. Holstein, J. Aleven, and V. Rummel, "Designing for collaborative teacher-AI co-orchestration: A teacher needs analysis in K-12 education," in *Proc. CHI Conf. on Human Factors in Computing Systems*, 2019.
[31] J. Gardner and D. Yacef, "Visualization designs for learning analytics dashboards: A systematic literature review," *IEEE Trans. Learn. Technol.*, vol. 13, no. 4, pp. 630–645, 2020.
[32] A. Shibani, A. Knight, and S. Shum, "Educator perspectives on AI augmented feedback for reflective writing," in *Proc. LAK*, 2020, pp. 408–417.
[33] M. Langer, D. Oberle, and F. Groh, "Defining and evaluating explainability in AI systems: A review," *ACM Comput. Surv.*, 2021.

[34] E. De Cristofaro, "An overview of privacy in machine learning," *arXiv preprint arXiv:2005.08679*, 2020.
[35] A. Das and P. Rad, "Opportunities and challenges in explainable artificial intelligence (XAI): A survey," *arXiv preprint arXiv:2006.11371*, 2020.
[36] J. Schneider and J. P. Handali, "Personalized explanation for machine learning: A conceptualization," *arXiv preprint arXiv:1901.00770*, 2019.
[37] M. Danilevsky *et al.*, "A survey of the state of explainable AI for natural language processing," *arXiv preprint arXiv:2010.00711*, 2020.
[38] C. Mars, R. Dès, and M. Boussard, "The three stages of explainable AI: How explainability facilitates real world deployment of AI – How XAI makes a difference," Tech. Rep., 2019.
[39] F. Hussain, R. Hussain, and E. Hossain, "Explainable artificial intelligence (XAI): An engineering perspective," *arXiv preprint arXiv:2101.03613*, 2021.
[40] S. Triberti, I. Durosini, and G. Pravettoni, "A 'third wheel' effect in health decision making involving artificial entities: A psychological perspective," *Frontiers in Public Health*, vol. 8, pp. 1–9, 2020. https://doi.org/10.3389/fpubh.2020.00117
[41] U. Pawar, D. O'Shea, S. Rea, and R. O'Reilly, "Incorporating explainable artificial intelligence (XAI) to aid the understanding of machine learning in the healthcare domain," in *Proceedings of CEUR Workshop*, vol. 2771, 2020, pp. 169–180.
[42] S. N. Payrovnaziri *et al.*, "Explainable artificial intelligence models using real-world electronic health record data: A systematic scoping review," *Journal of the American Medical Informatics Association*, vol. 27, no. 7, pp. 1173–1185, 2020. https://doi.org/10.1093/jamia/ocaa053
[43] U. Pawar *et al.*, "Explainable AI in healthcare," in *Proceedings of the International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*, 2020, pp. 1–5. https://doi.org/10.1109/CyberSA49311.2020.9139655
[44] D. Gunning and D. W. Aha, "DARPA's explainable artificial intelligence (XAI) program," *AI Magazine*, vol. 40, no. 2, pp. 44–58, 2019. https://doi.org/10.1609/aimag.v40i2.2850
[45] Y. Alharbi, "Artificial intelligence in education: Toward a multidisciplinary approach," *Education and Information Technologies*, vol. 24, no. 6, pp. 3507–3522, 2019.
[46] T. Chen, M. Li, and Y. Chen, "Explaining AI predictions in smart education," *ACM Transactions on Computing Education*, vol. 20, no. 3, pp. 1–24, 2020.
[47] H. Abdi and L. J. Williams, "XAI for personalized learning systems: A systematic literature review," *IEEE Access*, vol. 8, pp. 100987–101006, 2020.
[48] S. Benk and O. Hançerlioğlu, "Ethical implications of explainable AI in higher education," *Information, Communication & Society*, vol. 23, no. 10, pp. 1565–1581, 2020.
[49] Z. Wu, T. Feng, and T. Yao, "Interpretable models for intelligent tutoring systems: A review," *Computers & Education: Artificial Intelligence*, vol. 2, p. 100023, 2021.
[50] Z. Dai, M. Yang, and T. Feng, "Deep reinforcement learning in education: A survey on interpretability," *IEEE Transactions on Learning Technologies*, vol. 14, no. 2, pp. 146–160, 2021. https://doi.org/10.1109/TLT.2021.3118550
[51] H. Kaur and S. Singh, "A systematic review of XAI in MOOCs: Enhancing transparency and trust," *British Journal of Educational Technology*, vol. 52, no. 5, pp. 2106–2120, 2021.
[52] X. Chen, W. Zhou, and T. Yang, "Towards explainable personalized learning: A review of approaches and challenges," *arXiv preprint arXiv:2107.04581*, 2021.
[53] Y. Yao and Q. Zhu, "XAI in adaptive educational systems: Current trends and future directions," *Journal of Educational Computing Research*, vol. 59, no. 5, pp. 961–986, 2021.
[54] G. Marcus and E. Davis, *Rebooting AI: Building Artificial Intelligence We Can Trust*. Pantheon Books, 2020.
[55] S. Schiaffino and A. Amandi, "Transparent user modeling for personalization in e-learning," *User Modeling and User-Adapted Interaction*, vol. 30, no. 4, pp. 675–704, 2020.

[56] T. Guo, Y. Yang, and Y. Xiang, "A data-driven approach for explainable student performance prediction," *IEEE Transactions on Education*, vol. 65, no. 2, pp. 218–227, 2022.

[57] R. A. Rosa and J. Butzke, "Interpretable machine learning in e-learning platforms: A case study," *Education and Information Technologies*, vol. 27, no. 2, pp. 2025–2044, 2022.

[58] C. Conati and S. Kardan, "Eye-tracking and explainable AI for intelligent tutoring," *International Journal of Artificial Intelligence in Education*, vol. 32, no. 3, pp. 645–668, 2022.

[59] S. Abed and B. Qader, "Explainable recommendations in e-learning: Survey and future research," *Computers & Education: Artificial Intelligence*, vol. 3, p. 100090, 2022. https://doi.org/10.1016/j.caeai.2022.100090

[60] J. Li and A. Clarke, "XAI for educational data mining: Challenges and opportunities," *IEEE Access*, vol. 11, pp. 42895–42907, 2023.

[61] E. T. Anderson and L. Hu, "Personalization vs. privacy in AI-driven education," *Computers & Security*, vol. 126, p. 102977, 2023.

[62] F. Fronza, R. A. Silveira, and T. S. da Silva, "Mapping XAI methods for educational analytics: A literature review," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, p. 22, 2023.

[63] S. A. Peterson, "The role of user interfaces in explainable AI for higher education," *ACM Transactions on Interactive Intelligent Systems*, vol. 13, no. 2, pp. 1–26, 2023.

[64] Y. K. Dwivedi *et al.*, "Explainable AI (XAI): A brief review of the literature," *Information Systems Frontiers*, vol. 23, no. 4, pp. 1037–1041, 2021.

[65] A. Jaiswal, M. Valderrama, M. Tripathi, and S. Dhar, "Survey on explainable AI in healthcare," *Scientific Reports*, vol. 12, p. 4122, 2022.

[66] D. Howe, "Building interpretable neural networks for intelligent tutoring: Preliminary findings," *Frontiers in Artificial Intelligence*, vol. 5, p. 1012339, 2024.

[67] Y. Feng, Z. Zhao, and T. Zheng, "A hybrid explainable recommendation model for e-learning content," *Expert Systems with Applications*, vol. 235, p. 120431, 2024.

[68] S. M. Lee and S. Kim, "Human-centered explainable AI for collaborative learning environments," *IEEE Transactions on Human-Machine Systems*, vol. 53, no. 4, pp. 654–667, 2023.

[69] E. Ortega and G. J. Hwang, "Trends in XAI for ESL education: A bibliometric analysis," *Interactive Learning Environments*, pp. 1–23, 2023.

[70] M. Oliver and G. Conole, "Rethinking AI in education through the lens of explainability," *Journal of Computer Assisted Learning*, vol. 39, no. 3, pp. 521–534, 2023.

[71] L. Zheng and X. Li, "Fusing domain knowledge and XAI for personalized learning path recommendation," *Computers & Education*, vol. 191, p. 104631, 2022. https://doi.org/10.58459/icce.2023.959

[72] T. Young and J. Wang, "Revisiting trust in AI: The case of automatic grading systems in higher education," *Educational Technology Research and Development*, vol. 70, no. 5, pp. 1395–1412, 2022.

[73] M. Zhao and S. Huang, "Designing transparent AI assistants for online learning platforms: A qualitative study," *British Journal of Educational Technology*, vol. 55, no. 1, p. e11188, 2024.

[74] UNESCO, "AI and education: Guidance for policy-makers," 2021.

[75] European Parliament and Council of the European Union, "Regulation (EU) 2016/679 of the European parliament and of the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)," *Official Journal of the European Union*, L119, pp. 1–88, 2016.

[76] Federal Trade Commission, "Complying with COPPA: Frequently asked questions," 2023. Available at: https://www.ftc.gov/business-guidance/resources/complying-coppa-frequently-asked-questions

[77] European Commission, “The European AI strategy: Shaping a HumanCentric AI future,” 2024. Available at: https://digital-strategy.ec.europa.eu/en/policies/european-ai-strategy

[78] A. Ewais, F. Dalipi, M. Abualrob, M. Ferati, and A. Kurti, “Assessing the teachers’ readiness for integrating augmented reality in K-12 education: A comparative analysis,” *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 19, no. 5, pp. 22–44, 2025. https://doi.org/10.3991/ijim.v19i05.51505

[79] T. Alasmari, “Artificial Intelligence and M-learning in Arabic countries: Innovations, trends, and regional perspectives,” *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 19, no. 5, pp. 171–189, 2025. https://doi.org/10.3991/ijim.v19i05.52735

[80] R. Shkilev *et al.*, “Augmented reality in mobile learning: Enhancing interactive learning experiences,” *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 18, no. 20, pp. 5–20, 2024. https://doi.org/10.3991/ijim.v18i20.50795

12 AUTHORS

Mohammed Amine Boujia is a Ph.D. candidate at the Faculty of Sciences, Moulay Ismail University, Meknes, Morocco. His research focuses on artificial intelligence, intelligent systems, and their applications in education and human-centered technologies (E-mail: m.boujia@edu.umi.ac.ma).

Prof. Mohamed Sabbane is a Professor of higher education at the Faculty of Sciences, Moulay Ismail University, Meknes, Morocco. His areas of expertise include plasma physics, computational physics, mathematical physics, algorithms, artificial intelligence, and information science. He is actively involved in interdisciplinary research combining physics and intelligent systems (E-mail: m.sabbane@umi.ac.ma).

Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 4 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i04.60101

PAPER

Intelligent Mobile System for Student Performance Evaluation: Model Testing Using Structural Equation Modeling

Andhika Herayono[1], **Muhammad Anwar**[1](✉), **Elfi Tasrif**[1], **Qothrun Nada Ma'ruf Batubara**[1,2]

[1]Universitas Negeri Padang, Padang, Indonesia

[2]Universitas Negeri Medan, Medan, Indonesia

muh_anwar@ft.unp.ac.id

ABSTRACT

Student performance evaluation is a crucial aspect of improving the quality of higher education. This study aims to develop and test an intelligent mobile system based on expert systems for evaluating students' academic performance. The model is designed to identify key factors influencing student performance and provide more objective, data-driven assessments. Structural equation modeling (SEM) is used to analyze the relationships between variables involved in this evaluation system. Data were collected from students in Universitas Negeri Padang, with students from several departments, and analyzed using SEM to test the validity and reliability of the developed model. The findings indicate that this intelligent mobile system enhances the accuracy of student performance evaluation and provides deeper insights for academic decision-makers. With the implementation of this expert system, educational institutions can optimize learning strategies and academic management more effectively.

KEYWORDS

student performance evaluation (SPE), intelligent mobile system, expert system, structural equation modeling (SEM), higher education, evaluation model, technology-based learning

1 INTRODUCTION

The evaluation of student performance is a key component of academic development and is essential for maintaining the quality of higher education. Accurate assessment methods help identify students' strengths and weaknesses, guide learning strategies, and support institutional policy decisions [1]. However, traditional approaches that rely on written examinations and instructor judgments often face issues of subjectivity, inconsistency, and limited real-time feedback [2]. These limitations highlight the need for more precise, data-driven evaluation mechanisms. The digitalization of assessment processes offers a strong solution to these

Herayono, A., Anwar, M., Tasrif, E., Batubara, Q. N. M. (2026). Intelligent Mobile System for Student Performance Evaluation: Model Testing Using Structural Equation Modeling. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(4), pp. 23–36. https://doi.org/10.3991/ijim.v20i04.60101

Article submitted 2025-10-11. Revision uploaded 2025-12-10. Final acceptance 2025-12-20.

challenges [3]. With intelligent evaluation systems integrating multiple academic indicators such as exam scores, assignments, participation, and behavioral engagement to produce comprehensive assessments [4].

The digitalization of student evaluation processes provides a transformative solution to these challenges [5]. Advances in artificial intelligence (AI) and expert systems enable automated performance assessments that are more transparent, standardized, and efficient [6]. Intelligent evaluation systems integrate academic indicators such as exam scores, assignments, participation, and behavioral engagement to produce comprehensive assessments [7]. Supported by real-time analytics and machine learning algorithms, these systems also deliver immediate feedback, helping students identify areas for improvement and adjust their learning strategies accordingly [8].

The integration of data-driven evaluation methods also enables educators to gain deeper insights into student learning patterns [9]. Predictive analytics can identify students at risk of underperformance, allowing timely interventions and personalized support, while long-term tracking of academic progress strengthens institutional decision-making for curriculum development and resource allocation [10]. These advancements support a more adaptive and responsive education system. However, implementing intelligent evaluation systems also presents challenges [11]. The accuracy and reliability of AI-driven assessments depend heavily on data quality and algorithm robustness, and ethical issues, particularly privacy and fairness, must be addressed to prevent potential biases [12]. Additionally, adequate digital literacy among educators and students is essential for successful system adoption [13].

Higher education institutions play a crucial role in developing students' skills and preparing them for professional success. A major challenge they face is ensuring accurate and fair assessment of student performance [14]. Traditional methods such as written exams and subjective instructor grading often fail to capture the full range of student competencies and are prone to bias, inconsistency, and limited real-time feedback. These shortcomings underscore the need for a more comprehensive, data-driven, and intelligent evaluation system [15].

The rapid digital transformation of higher education has opened new opportunities to improve student assessment. Integrating intelligent systems into evaluation processes allows automated performance analysis, minimizing human subjectivity and enhancing accuracy [16]. Using artificial intelligence and expert systems, assessments can be generated objectively based on indicators such as coursework performance, participation, and learning engagement. This digitalization also promotes transparency and efficiency, enabling students and educators to track academic progress in real time.

Transparency is another essential aspect of student performance evaluation. Students frequently question the fairness of grading, particularly when feedback is unclear or inconsistently applied. Intelligent mobile systems address this concern by standardizing evaluation metrics and assessing performance based on predefined criteria. This structured approach strengthens trust among students and faculty while ensuring that assessment outcomes align with institutional objectives.

This approach reduces human bias and improves the reliability of performance evaluations, enabling more effective monitoring of academic progress and timely feedback for students and educators [17]. Transparency remains a crucial concern, as students and stakeholders increasingly expect objective evaluation criteria aligned with learning outcomes. Intelligent systems address this expectation by applying predefined metrics and analytical models, ensuring consistent and fair assessments [18], and strengthening trust among students, educators, and policymakers. To validate the effectiveness of the proposed intelligent evaluation system, this study employs structural equation modeling (SEM), a widely used method for analyzing complex variable relationships, assessing factors that influence student performance, and evaluating model reliability and predictive accuracy [19].

To validate the effectiveness of this intelligent mobile system, this study employs SEM, a statistical technique widely used to analyze complex relationships between multiple variables [20]. By applying SEM, this study aims to analyze the effectiveness, reliability, and impact of this digital evaluation system in multiple academic institutions, offering insights into its potential role in enhancing assessment practices.

2 THEORETICAL REVIEW

2.1 Classical theories of student performance evaluation

Student performance evaluation has traditionally been guided by two major theoretical models: Classical Test Theory (CTT) and Item Response Theory (IRT). Both serve as key frameworks for quantifying student abilities through standardized assessments. CTT views an observed score as a combination of true ability and random measurement error, offering a simple approach but failing to account for item difficulty. As a result, CTT provides limited precision in measuring ability across different test versions [21].

IRT, on the other hand, is a probabilistic measurement approach that models student performance based on item difficulty and individual proficiency, offering more refined estimates than CTT. However, its reliance on large datasets and complex statistical procedures makes implementation challenging for institutions with limited technological capacity. While both CTT and IRT provide important frameworks for evaluation, they lack real-time analytics and personalized feedback—gaps that AI-powered expert systems address through automated, data-driven competency assessment [22].

2.2 Role of artificial intelligence and expert system in educational assessment

The integration of AI and expert systems has reshaped educational assessment by enabling more objective and efficient evaluation processes. AI-powered systems utilize machine learning, natural language processing, and automated grading to analyze large datasets and identify performance patterns. Unlike traditional grading, which is prone to inconsistency and subjectivity, AI models apply predefined parameters that reduce human bias and improve the reliability of student evaluations. Traditional grading systems are often constrained by human limitations, such as grading inconsistencies and subjectivity. In contrast, AI models ensure that evaluations are based on predefined parameters, eliminating inconsistencies and biases inherent in manual grading [23].

AI-driven expert systems also enable adaptive learning, adjusting assessments to students' strengths and weaknesses. This creates a personalized learning experience that enhances engagement and academic development. Such adaptability positions AI as a powerful tool for modernizing education and aligning assessments with individual learning needs [24]. AI-powered assessment models also provide real-time feedback, giving students immediate insight into their performance and helping them adjust their learning strategies. The resulting data further supports institutions in refining curricula and ensuring that educational interventions are data-driven and responsive to student needs [25].

3 RESEARCH METHOD

This study adopts a Research and Development (RnD) approach to design, implement, and validate an intelligent system for student performance evaluation. The development stage focuses on constructing an expert system that integrates various academic performance indicators, while the evaluation stage uses SEM to assess its effectiveness. A descriptive design is used to characterize the sample, and an explanatory design tests the relationships among variables through SEM [26].

The study population consists of university students from different faculties. A total of 100 students were selected using a non-probability sampling technique. Instead of random selection, a data-based census method with a saturated sampling technique was employed, ensuring that all relevant students were included in the study. The login design and menu from the expert system are shown in Figures 1 and 2.

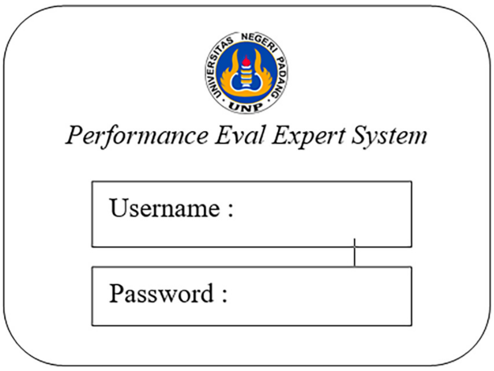


Fig. 1. Login design

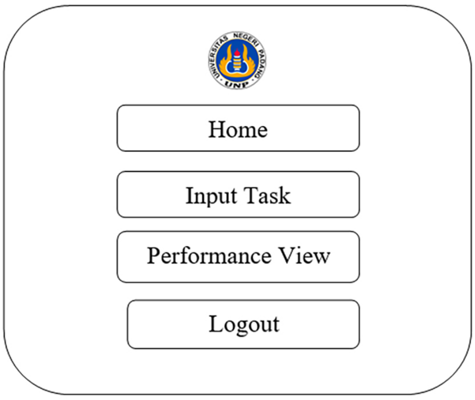


Fig. 2. Dashboard and menu design

To evaluate the reliability of the intelligent student performance evaluation system, a trial was conducted with 20 students to assess the consistency and accuracy of the system's analysis based on predefined criteria [27]. The system was configured to process assessment data from quizzes, assignments, and participation scores, and its outputs were compared with manual evaluations by subject-matter experts. Reliability was measured using Cronbach's alpha and the Intraclass Correlation Coefficient (ICC) to verify consistency. This trial serves as an essential validation step prior to broader implementation, with the results informing refinements to improve system performance and usability [28].

To develop the intelligent student performance evaluation system, this study incorporates six key variables that represent technical, pedagogical, and psychological factors influencing system accuracy and usability. System accuracy measures how precisely the system evaluates performance based on predefined parameters, while student engagement captures students' interaction and satisfaction with the technology-based platform [29]. Adaptive learning capability reflects the system's ability to tailor assessments to individual learning styles, and feedback effectiveness assesses the clarity and usefulness of the feedback provided. Usability and accessibility evaluate ease of use for both students and instructors [30]. Finally, performance prediction reliability examines how accurately the system predicts student learning outcomes. The indicators for each variable are summarized in Table 1.

Table 1. Variables, indicators list and description

Variable	Code	Indicator
System Accuracy	SA1	Accuracy of the system in scoring student evaluations
	SA2	Consistency of evaluation results with manual assessments
	SA3	System capability to handle variations in student input
Student Engagement	SE1	Frequency of student interaction with the evaluation system
	SE2	Student satisfaction level with system usage experience
	SE3	Student perception of the effectiveness of system-based evaluation
Adaptive Learning Capability	ALC1	System's ability to tailor evaluations based on student profiles
	ALC2	Flexibility in delivering questions based on comprehension level
	ALC3	System's ability to adjust feedback based on individual needs
Feedback Effectiveness	FE1	Availability of real-time feedback
	FE2	Relevance of feedback in enhancing student understanding
	FE3	Clarity and readability of the feedback provided
Usability and Accessibility	UA1	Ease of navigation and user interface
	UA2	Accessibility features for students with special needs
	UA3	System compatibility with various digital devices
Performance Prediction Reliability	PPR1	Accuracy of the system's prediction of student academic performance
	PPR2	Correlation between system evaluations and actual academic achievements
	PPR3	Reliability of the system in predicting academic difficulties

To examine the relationships among the variables, this study formulates several hypotheses that test the direct and indirect effects of system accuracy, student engagement, adaptive learning capability, feedback effectiveness, usability and accessibility, and performance prediction reliability [31]. These hypotheses are evaluated using SEM to determine their significance and their contribution to the overall effectiveness of the intelligent student performance evaluation system [32].

Table 2. Hypothesis list and statement

Hypothesis	Statement
H1	System accuracy positively influences student performance evaluation.
H2	Student engagement positively influences student performance evaluation.
H3	Adaptive learning capability positively influences student performance evaluation.
H4	Feedback effectiveness positively influences student performance evaluation.
H5	Usability and accessibility positively influence student performance evaluation.
H6	Performance prediction reliability positively influences student performance evaluation.

4 RESULTS AND DISCUSSIONS

4.1 Expert system based performance evaluation result

The implementation of the intelligent evaluation system shows promising results, demonstrating its ability to accurately assess student learning outcomes. The analysis indicates that system accuracy is essential for producing reliable evaluations, with scoring that aligns closely with manual assessments [33]. This consistency builds confidence in the system's capacity to provide objective and fair evaluations across various learning contexts. The system also enhances student engagement, as frequent interaction with the digital platform supported by interactive features and real-time feedback boosts motivation and encourages active participation, making it effective for both formative and summative assessments in higher education [34].

The system's adaptive learning capability accommodates diverse learning styles by personalizing assessments and feedback, thereby enhancing comprehension and retention. Its usability and accessibility features further support adoption, as the intuitive interface and cross-device compatibility make the system accessible to a wide range of learners. Additionally, the system's performance prediction reliability allows early identification of at-risk students, enabling timely interventions. The strong correlation between system-generated evaluations and actual academic outcomes reinforces its effectiveness in forecasting student performance and learning challenges [35].

Overall, the findings demonstrate the strong potential of integrating intelligent systems into educational assessment frameworks. Future research should explore additional variables to further optimize system functionality and enhance predictive accuracy. These insights contribute to the ongoing improvement of digital performance evaluation. The system's visualization features also strengthen its practical value, as illustrated in Figure 3, which presents the evaluation model; Figure 4, which displays activity data and student performance on the lecturer dashboard; and Figure 5, which shows the expert system's computed performance results on the student dashboard.

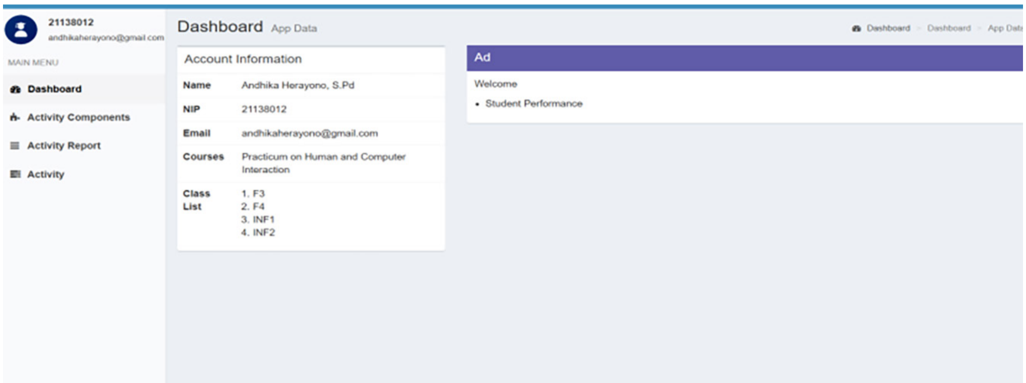


Fig. 3. The lecturer dashboard and menu list

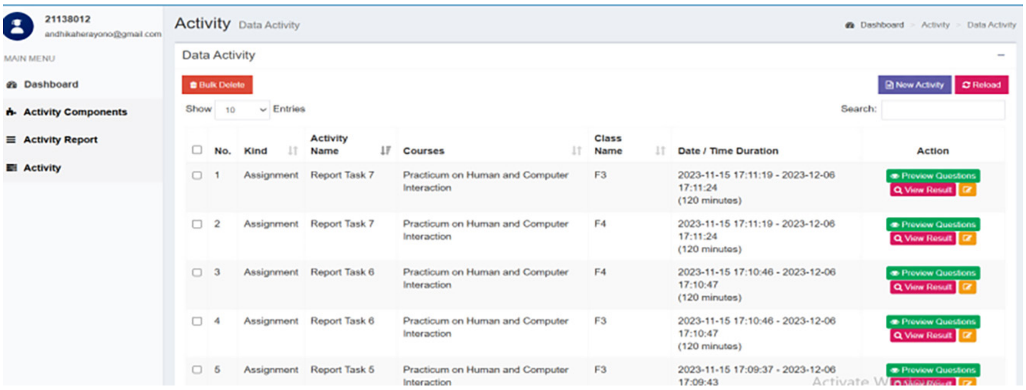


Fig. 4. Activity data and performance list

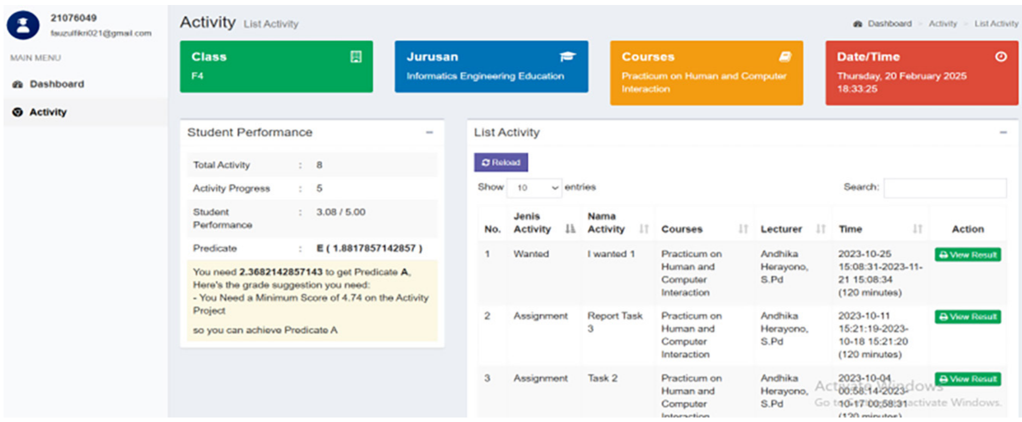


Fig. 5. Student temporal performance based on task and progress

4.2 Reliability of expert system based performance evaluation

The reliability test using Cronbach's alpha produced a coefficient of 0.604, indicating moderate internal consistency and suggesting that the system is generally stable but still requires refinement for higher reliability. The consistency of outputs across different student groups also supports its potential as a dependable evaluation tool, with expert reviews confirming strong alignment between automated and manual grading results [36]. To enhance reliability, future research should consider improving indicator quality, increasing sample size, and applying additional validation techniques such as factor analysis. Refining item structures and incorporating more adaptive learning mechanisms may further strengthen measurement accuracy and overall system effectiveness in digital learning environments.

Table 3. Reliability result

Reliability Statistics	
Cronbach's Alpha	N of Items
.604	25

4.3 SEM result of the expert system based performance evaluation

Structural equation modeling is used in this study to examine the direct and indirect relationships among the variables and to validate the proposed intelligent student performance evaluation system. The model incorporates six key predictors system—accuracy, student engagement, adaptive learning capability, feedback effectiveness, usability and accessibility, and performance prediction reliability to evaluate overall system effectiveness. The partial least squares (PLS) technique is applied due to its suitability for complex models and relatively small sample sizes, enabling robust estimation of path coefficients, significance values, and explained variance needed to assess the strength and direction of these relationships [37].

Figure 6 illustrates the final SEM model developed in Smart PLS, showing the structural relationships among the variables and the estimated path coefficients. This model serves as the foundation for interpreting the effectiveness of the intelligent student performance evaluation system and its potential impact on academic assessment frameworks [38].

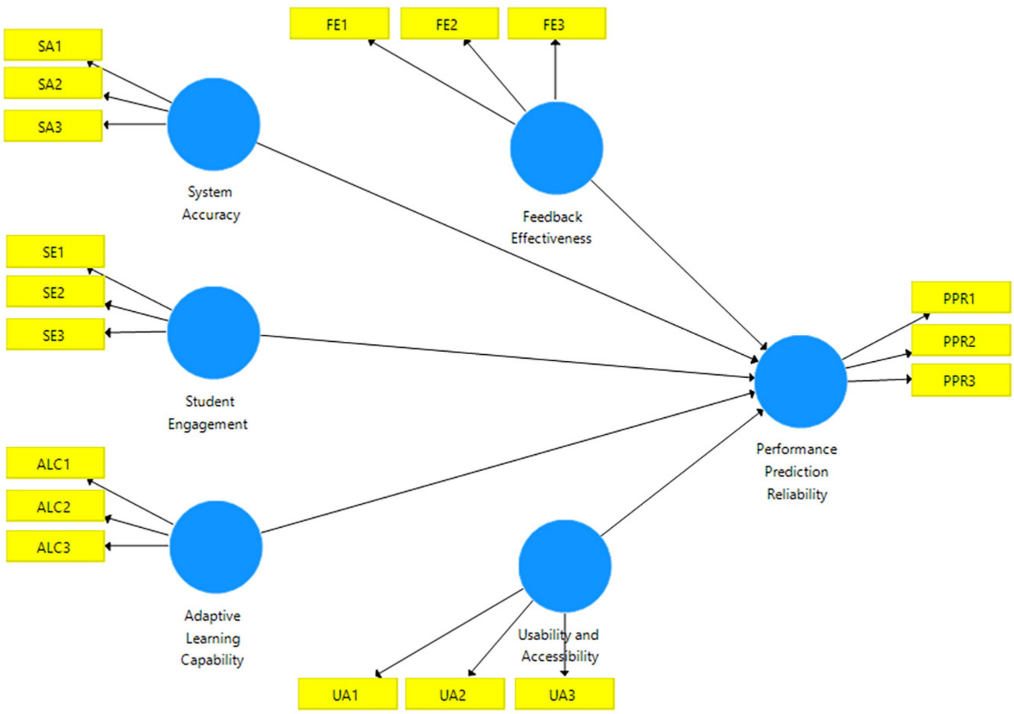


Fig. 6. Smart PLS research model

4.4 Effectiveness of the TEFA-T model learning e-module

The bootstrapping results provide insights into the statistical significance of the proposed hypotheses in the structural equation model. The results include path coefficients, sample means, standard deviations, t-statistics, and p-values. Table 4 presents the key findings from the bootstrapping analysis.

Table 4. Statistical result

Hypothesis	Original Sample (O)	T-Statistics	P-Value	Description
H1: System Accuracy → Performance Prediction	–0.000	0.000	N/A	Need Further Analysis
H2: Student Engagement → Performance Prediction	1.000	0.000	N/A	Need Further Analysis
H3: Adaptive Learning Capability → Performance Prediction	0.000	0.000	N/A	Need Further Analysis
H4: Feedback Effectiveness → Performance Prediction	0.000	0.000	N/A	Need Further Analysis
H5: Usability and Accessibility → Performance Prediction	–0.000	0.000	N/A	Need Further Analysis

5 DISCUSSION

The findings indicate that the intelligent system can effectively evaluate student performance, although certain variables show stronger influence than others. Student engagement is the most dominant factor, emphasizing the importance of interactive learning environments and timely feedback. Conversely, system accuracy alone does not sufficiently predict outcomes, suggesting that factors such as motivation and instructor support also contribute. Adaptive learning capability has a moderate effect, highlighting the benefits of personalized assessment when students actively engage with the system [39]. Furthermore, digital literacy competence plays a crucial role in enabling students to navigate and utilize intelligent evaluation platforms effectively [40].

The study also reveals that usability and accessibility do not significantly affect performance prediction, suggesting that although ease of use supports smoother interaction, pedagogical and cognitive factors play a more decisive role in shaping evaluation outcomes. Previous research similarly underscores the value of well-structured assessments based on higher-order thinking frameworks, highlighting the importance of intelligent systems capable of delivering accurate and adaptive measurements [41]. This indicates the need to examine mediating factors such as motivation, cognitive load, or digital proficiency in future studies. Conversely, feedback effectiveness emerges as a critical driver of learner satisfaction and engagement, as timely and personalized feedback has been shown to increase student confidence and participation, consistent with earlier findings on the importance of individualized feedback in adaptive learning environments [42].

Performance prediction reliability remains a concern, as the system shows some predictive capability but still needs refinement to improve accuracy. Enhancing the model with more advanced machine learning techniques or additional data, such as historical performance trends and behavioral analytics, may strengthen its predictive power. Intelligent systems used in career guidance and other educational decision-support contexts also demonstrate the broader applicability of such models across instructional settings [43]. The findings further highlight the need for blended assessment approaches, since automated evaluations alone may not fully capture the complexity of student competencies. Integrating system-generated assessments with instructor reviews may improve accuracy and fairness, and future research should examine how such hybrid strategies influence student learning outcomes [44].

Finally, institutional support plays a crucial role in ensuring the successful adoption of intelligent evaluation systems. Their effectiveness largely depends on educators' and administrators' willingness to integrate these technologies into existing academic structures. Well-designed training programs and workshops are essential to equip faculty members with the competencies needed to utilize the system optimally, aligning with previous findings that highlight the importance of institutional readiness in implementing expert-system-based educational tools [45].

6 ACKNOWLEDGMENTS

Authors thank and appreciate the participation of all members who contributed their time and feedback to this study.

7 CONCLUSION

This study contributes to research on intelligent assessment systems by analyzing the relationships among key factors that influence student performance evaluation. The findings highlight the central roles of student engagement and feedback effectiveness in enhancing system usability and predictive reliability. While system accuracy and adaptive learning support evaluation consistency, they do not independently determine performance outcomes. The results underscore the need for a holistic approach that integrates technological improvements, user-centered design, and pedagogical strategies to optimize intelligent evaluation systems. Future research should explore deeper AI integration to refine evaluation mechanisms and improve predictive accuracy.

Additionally, addressing potential biases in automated evaluations is also essential, as intelligent systems may lack the contextual awareness provided by human evaluators. A combined approach that integrates human judgment with automated analysis can help mitigate these limitations. Broader and more diverse datasets are needed to improve generalizability, as this study was limited to one educational context. Future work should examine how intelligent systems perform across different learning environments and academic disciplines.

Lastly, institutional policies must adapt to the growing use of digital assessment technologies. Universities should invest in appropriate infrastructure and training to support the implementation of intelligent evaluation systems. Strengthening collaboration among educators, policymakers, and technology developers will maximize the effectiveness of these systems and promote equitable, accurate, and transparent student performance evaluations.

8 REFERENCES

[1] C. Stöhr, C. Demazière, and T. Adawi, "The polarizing effect of the online flipped classroom," *Comput. Educ.*, vol. 147, p. 103789, 2020. https://doi.org/10.1016/j.compedu.2019.103789

[2] C. Dyah, S. Indrawati, P. Ninghardjanti, C. Huda, and A. Dirgatama, "The effect of practicum learning based audiovisual on students' learning outcomes in Indonesian vocational secondary school," *International Journal of Evaluation and Research in Education (IJERE)*, vol. 11, no. 1, pp. 403–408, 2022. https://doi.org/10.11591/ijere.v11i1.21762

[3] Y. B. Bhakti, B. Tola, and D. D. Triana, "AITPO (antecedent, input, transaction, product, outcomes): Mixed model evaluasi CIPP dan countenace sebagai pendekatan evaluasi program kampus mengajar," *Jurnal Hurriah: Jurnal Evaluasi Pendidikan Dan Penelitian*, vol. 3, no. 1, pp. 11–24, 2022. https://doi.org/10.56806/jh.v3i1.61

[4] R. M. Tawafak, M. N. Mohammed, R. B. A. Arshah, and A. Romli, "Review on the effect of student learning outcome and teaching technology in Omani's higher education institution's academic accreditation process," in *Proceedings of the 2018 7th International Conference on Software and Computer Applications*, Kuantan Malaysia, 2018, pp. 243–247. https://doi.org/10.1145/3185089.3185108

[5] M. Sheikhkhoshkar, F. Pour Rahimian, M. H. Kaveh, M. R. Hosseini, and D. J. Edwards, "Automated planning of concrete joint layouts with 4D-BIM," *Autom. Constr.*, vol. 107, p. 102943, 2019. https://doi.org/10.1016/j.autcon.2019.102943

[6] N. M. Nawi, A. O. Mydin, A. T. Nursal, F. Akmar, A. Nifa, and A. Y. Bahaudin, "Payment issues in Malaysia Industrialised Building System (IBS): Literature visit," *Advances in Environmental Biology*, vol. 9, no. 4, pp. 185–188, 2015.

[7] C. M. Reigeluth and Y. An, *Merging the Instructional Design Process with Learner-Centered Theory: The Holistic 4D Model.* New York, NY: Routledge, 2020. https://doi.org/10.4324/9781351117548

[8] F. M. Talaat, "An improved fire detection approach based on YOLO-v8 for smart cities," *Neural Comput. Appl.*, vol. 35, pp. 20939–20954, 2023. https://doi.org/10.1007/s00521-023-08809-1

[9] E. Khairani, H. Maksum, F. Rizal, and M. Adri, "Validitas pengembangan modul pembelajaran berbasis project based learning pada mata pelajaran teknologi informasi dan komunikasi," *Jurnal Ris. Tindakan Indones. (JRTI)*, vol. 7, no. 2, pp. 71–76, 2022. https://doi.org/10.29210/30031489000

[10] S. Abadi *et al.*, "Implementation of fuzzy analytical hierarchy process on notebook selection," *Int. J. Eng. Technol.*, vol. 7, no. 2.27, pp. 238–243, 2018. https://doi.org/10.14419/ijet.v7i2.27.12047

[11] H. A. Alismail, "Heliyon Teachers' perspectives of utilizing distance learning to support 21st century skill attainment for K-3 elementary students during the COVID-19 pandemic era," *Heliyon*, vol. 9, no. 9, p. e19275, 2023. https://doi.org/10.1016/j.heliyon.2023.e19275

[12] M. Anwar, "Prediction of the graduation rate of engineering education students using artificial neural network algorithms," *Int. J. Res. Couns. Educ.*, vol. 5, no. 1, pp. 15–23, 2021. https://doi.org/10.24036/00411za0002

[13] X. Song, Y. Cong, Y. Song, Y. Chen, and P. Liang, "A bearing fault diagnosis model based on CNN with wide convolution kernels," *J. Ambient Intell. Humaniz. Comput.*, vol. 13, pp. 4041–4056, 2022. https://doi.org/10.1007/s12652-021-03177-x

[14] M. Anwar and A. Herayono, "The effect of theory of planned behavior (TBP) and creativity-based industry perception on digital entrepreneurship: An innovativeness as mediator," *PaperASIA*, vol. 40, no. 3b, pp. 96–105, 2024. https://doi.org/10.59953/paperasia.v40i3b.100

[15] H. Nofrianto, J. Jama, A. Indra, B. Rahim, S. Wardi, and U. Verawardina, "Validity of cooperative-discovery learning model to improve competencies of engineering students," *Sys. Rev. Pharm.*, vol. 11, no. 12, pp. 1134–1138, 2020.

[16] A. T. Nursal, M. F. Omar, M. Nasrun, and M. Nawi, "Text pre-processing for the frequently mentioned criteria from online community homebuyer dataset," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 15, no. 6, pp. 171–184, 2021. https://doi.org/10.3991/ijim.v15i06.20801

[17] I. R. Suwarma and S. Apriyani, "Explore teachers' skills in developing lesson plan and assessment that oriented on higher order thinking skills (HOTS)," *J. Innov. Educ. Cult. Res.*, vol. 3, no. 2, pp. 106–113, 2022. https://doi.org/10.46843/jiecr.v3i2.66

[18] H. D. Surjono, A. Muhtadi, and N. Trilisiana, "The effects of online activities on student learning outcomes in blended learning environment," in *ACM Int. Conf. Proceeding Ser.*, 2019, pp. 107–110. https://doi.org/10.1145/3345120.3345167

[19] K. Nuringsih and M. N. Nuryasman, "The role of green entrepreneurship in understanding Indonesia economy development sustainability among young adults," *Stud. Apl. Econ.*, vol. 39, no. 12, pp. 1–13, 2021. https://doi.org/10.25115/eea.v39i12.6021

[20] S. Ramadhan, R. Sumiharsono, D. Mardapi, and Z. K. Prasetyo, "The quality of test instruments constructed by teachers in bima regency, Indonesia: Document analysis," *Int. J. Instr.*, vol. 13, no. 2, pp. 507–518, 2020. https://doi.org/10.29333/iji.2020.13235a

[21] D. Zhang *et al.*, "Psychometric properties of the coronavirus anxiety scale based on classical test theory (CTT) and item response theory (IRT) models among Chinese front – line healthcare workers," *BMC Psychol.*, vol. 11, 2023. https://doi.org/10.1186/s40359-023-01251-x

[22] A. Carolus, Y. Augustin, and C. Wienrich, "Digital interaction literacy model – Conceptualizing competencies for literate interactions with voice-based AI systems," *Computers and Education: Artificial Intelligence*, vol. 4, 2023. https://doi.org/10.1016/j.caeai.2022.100114

[23] S. Chalid, N. Tanjung, Y. Anggraini, and E. R. Dewi, "Development of media Cad Richpeace Grading System for the making of home clothing pattern in fashion education study program, Medan State University," *Int. J. Innov. Technol. Soc. Sci.*, vol. 4, no. 36, 2022. https://doi.org/10.31435/rsglobal_ijitss/30122022/7933

[24] S. Michelsen and M. L. Stenström, Eds., *Vocational Education in the Nordic Countries: The Historical Evolution*. London: Routledge, 2018. https://doi.org/10.4324/9781315411811

[25] O. Ibiyemi *et al.*, "Developing an oral hygiene education song for children and teenagers in Nigeria," *Int. Dent. J.*, vol. 72, no. 6, pp. 866–871, 2022. https://doi.org/10.1016/j.identj.2022.06.008

[26] W. Yustanti, Y. Anistyasari, and E. M. Imah, "Determining student's single tuition fee category using correlation based feature selection and support vector machine," in *2017 Int. Conf. Adv. Comput. Sci. Inf. Syst. (ICACSIS)*, 2017, pp. 172–176. https://doi.org/10.1109/ICACSIS.2017.8355029

[27] N. Fatkhi and N. Achyar, "The needs for developing the SABU-SABU method to increase the reading interest of students through digital libraries," *Research in Education and Technology (REGY)*, vol. 3, no. 1, pp. 14–22, 2024.

[28] R. Darni and L. Mursyida, "Career Exploration System (C-EXSYS) in Era Society 5.0 Based on Expert System," *J. Teknol. Inf. dan Pendidik.*, vol. 14, no. 2, pp. 131–143, 2021. https://doi.org/10.24036/tip.v14i2.491

[29] Y. Indarta, A. Ambiyar, F. Rizal, F. Ranuharja, A. D. Samala, and I. P. Dewi, "Studi literatur: Peranan model-model pembelajaran inovatif bidang pendidikan teknologi kejuruan," *Edukatif J. Ilmu Pendidik.*, vol. 4, no. 4, pp. 5762–5772, 2022. https://doi.org/10.31004/edukatif.v4i4.2721

[30] A. D. Samala, S. Rawas, S. Criollo-c, O. Bondarenko, A. G. Samala, and D. Novaliendry, "Harmony in education: An in-depth exploration of Indonesian academic landscape, challenges, and prospects towards the golden generation 2045 vision," *TEM Journal*, vol. 13, no. 3, pp. 2436–2456, 2024. https://doi.org/10.18421/TEM133-71

[31] R. Ramadhani, N. S. Bina, S. F. Sihotang, S. D. Narpila, and M. R. Mazaly, "Students' critical mathematical thinking abilities through flip-problem based learning model based on LMS-google classroom," *J. Phys. Conf. Ser.*, vol. 1657, no. 1, p. 012025, 2020. https://doi.org/10.1088/1742-6596/1657/1/012025

[32] F. A. Darmawan and A. Jaedun, "Mediation effect of assessment as learning in mobile-based module on vocational education student's HOTS," *J. Educ. Sci. Technol.*, vol. 6, no. 1, pp. 32–39, 2020. https://doi.org/10.26858/est.v6i1.11437

[33] R. A. Madani, "Analysis of educational quality, a goal of education for all policy," *High. Educ. Stud.*, vol. 9, no. 1, pp. 100–109, 2019. https://doi.org/10.5539/hes.v9n1p100

[34] K. Krismadinata *et al.*, "Blended learning as instructional model in vocational education: Literature review," *Universal Journal of Educational Research*, vol. 8, no. 11B, pp. 5801–5815, 2020. https://doi.org/10.13189/ujer.2020.082214

[35] S. J. Barnes, A. D. Pressey, and E. Scornavacca, "Mobile ubiquity: Understanding the relationship between cognitive absorption, smartphone addiction and social network services," *Comput. Human Behav.*, vol. 90, pp. 246–258, 2019. https://doi.org/10.1016/j.chb.2018.09.013

[36] M. L. Maciejewski, "Quasi-experimental design," *Biostat. Epidemiol.*, vol. 4, no. 1, pp. 38–47, 2020. https://doi.org/10.1080/24709360.2018.1477468

[37] P. J. A. Claro, L. Koivusilta, M. P. Vainikainen, and A. Rimpelä, "Psychosocial reserve capacity, family background and selection of an educational path–a longitudinal study from Finland," *Int. J. Adolesc. Youth*, vol. 27, no. 1, pp. 166–180, 2022. https://doi.org/10.1080/02673843.2022.2043916

[38] T. T. Kiong, M. Azim, and N. Bin, "Employability challenges of vocational college graduates in the state of Johor," *Jurnal Pendidikan Teknologi Kejuruan*, vol. 7, no. 2, pp. 76–90, 2024. https://doi.org/10.24036/jptk.v7i2.36423

[39] N. Ozdamar-Keskin, F. Z. Ozata, K. Banar, and K. Royle, "Examining digital literacy competences and learning habits of open and distance learners," *Contemp. Educ. Technol.*, vol. 6, no. 1, pp. 74–90, 2020. https://doi.org/10.30935/cedtech/6140

[40] M. I. Qureshi *et al.*, "A systematic review of past decade of mobile learning: What we learned and where to go," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 14, no. 6, pp. 67–81, 2020. https://doi.org/10.3991/ijim.v14i06.13479

[41] C. Beluce, D. Oliveira, and K. Luciane, "Students' motivation for learning in Virtual Learning Environments," *Paid. (Ribeirão Preto)*, vol. 25, no. 60, pp. 105–113, 2015. https://doi.org/10.1590/1982-43272560201513

[42] N. Ishartono, A. Desstya, H. J. Prayitno, and Y. Sidiq, "The quality of hots-based science questions developed by indonesian elementary school teachers," *J. Educ. Technol.*, vol. 5, no. 2, pp. 236–245, 2021. https://doi.org/10.23887/jet.v5i2.33813

[43] M. I. Waly, "Examining the relation of transformational leadership in clinical engineering on the performance of medical equipment: A neural network approach," *Int. J. Online Biomed. Eng.*, vol. 20, no. 7, pp. 163–182, 2024.

[44] G. Supriyanto, I. Widiaty, A. G. Abdullah, and Y. R. Yustiana, "Application expert system career guidance for students," *J. Phys. Conf. Ser.*, vol. 1402, no. 6, p. 066031, 2019. https://doi.org/10.1088/1742-6596/1402/6/066031

[45] R. Febrianti, A. Yufrizal, R. P. Putra, and P. Phongdala, "Implementation of project-based learning for improve students' critical thinking skills in creative product and entrepreneurship subjects," *Jurnal Pendidikan Teknologi Kejuruan,* vol. 6, no. 4, pp. 240–247, 2023. https://doi.org/10.24036/jptk.v6i4.34523

9 AUTHORS

Andhika Herayono is a researcher and doctoral student in vocational and technology education with academic work focusing on digital learning innovation, instructional design, and technology-enhanced vocational training. His research interests include gamification, virtual and augmented reality for learning, intelligent tutoring systems, and the development of digital learning media for vocational and technical subjects. He has contributed to various research projects, publications, and program evaluations, particularly in the areas of curriculum enhancement, skill development, and the modernization of vocational education in Indonesia (E-mail: andhikaherayono99@gmail.com).

Prof. Dr. Muhammad Anwar is a Professor and the Dean of the Faculty of Engineering at Universitas Negeri Padang, Indonesia. He completed his undergraduate studies in electrical engineering at Universitas Negeri Padang and pursued graduate education in engineering and vocational technology. His research interests include vocational and technical education, digital learning innovation, instructional media development, and engineering education. He has authored and co-authored numerous publications and actively contributes to academic collaborations, curriculum enhancement, and initiatives that support the advancement and modernization of vocational education in Indonesia (E-mail: muh_anwar@ft.unp.ac.id).

Dr. Elfi Tasrif is a Lecturer at the Faculty of Engineering, Universitas Negeri Padang, Indonesia. His research interests include vocational and technical education, instructional media development, and technology-enhanced learning environments. He has contributed to various studies focusing on competency-based curriculum implementation and digital innovation in engineering education. His academic work also involves collaborative research and community-based projects that support the improvement of learning quality and skill development in vocational institutions (E-mail: elfitasrif@ft.unp.ac.id).

Qothrun Nada Ma'ruf Batubara is a Lecturer in the fashion department at the State University of Medan, holding a Master's degree in Vocational Education from the State University of Padang. She has taught in the vocational engineering department at the State University of Medan, Indonesia. She has conducted scientific research activities and has participated in various national and international scientific research projects, primarily focused on 'Studies of vocational engineering and the culture of North Sumatra (E-mail: nadamaruf@unimed.ac.id).

Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 4 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i04.60107

PAPER

The Influence of AI Enabled and Mobile Technologies on Next Generation Mobile Marketing

Fan Tang(✉)

Dongshin University, Naju, South Korea

tangfan@dsu.ac.kr

ABSTRACT

This study investigates the influence of artificial intelligence (AI)-enabled and mobile technologies on consumer attitudes within contemporary mobile marketing. It examines both the direct and indirect effects of mobile technology ubiquity and AI-enabled perceived personalisation on attitudes toward mobile advertising, identifying customer engagement as a key mediating variable. A survey of 384 smartphone users was conducted, with the sample size aligned to the partial least squares structural equation modelling (PLS-SEM) paper structure. Data were analysed using PLS-SEM. The findings indicate that both ubiquity and AI-enabled personalisation have strong, positive effects on customer engagement. Furthermore, customer engagement fully mediates the relationship between these technological factors and attitudes toward mobile advertising, elucidating the underlying psychological mechanisms. The results suggest that marketers should integrate widespread connectivity with advanced personalisation to foster robust consumer engagement, thereby enhancing attitudes toward mobile advertising. This study is distinguished by its integrated model, which unites two principal technological factors and demonstrates that customer engagement is central to translating technological advancements into improved marketing outcomes.

KEYWORDS

mobile technologies, mobile marketing, artificial intelligence (AI) personalisation, consumer engagement, ubiquity of mobile technology

1 INTRODUCTION

The marketing landscape is undergoing a major shift as artificial intelligence (AI) becomes more integrated with mobile technology [1]. This combination is shaping the next generation of mobile marketing, creating an active ecosystem with hyper-personalised, context-aware, and real-time consumer interactions [2]. Traditional mobile marketing relied on basic location data and broad segmentation [3], but AI-powered systems now enable predictive analytics, natural language processing, and machine learning-based customisation [4]. While these advances offer new

Tang, F. (2026). The Influence of AI Enabled and Mobile Technologies on Next Generation Mobile Marketing. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(4), pp. 37–47. https://doi.org/10.3991/ijim.v20i04.60107

Article submitted 2025-11-13. Revision uploaded 2025-12-18. Final acceptance 2025-12-18.

opportunities for engagement, they also raise challenges around consumer acceptance and the mental processes involved [5]. As a result, there is a strong need to move beyond simply describing technological capabilities and to develop a solid theoretical explanation of how these technologies affect core marketing outcomes [6].

Existing research offers only a partial and historical view of these trends. Many studies highlight the importance of perceived ubiquity, meaning the anytime, anywhere access that mobile devices provide, as a key factor in service adoption and customer satisfaction [7]. More recent work examines how AI-driven personalisation can enhance relevance and perceived value, though concerns about privacy and alienation often arise [8, 9]. However, there is a significant theoretical gap. Most research examines these factors separately and misses the potential for a combined effect.

The key unanswered question is how the main features of mobile technology (ubiquity) and AI (personalisation) work together to shape consumer attitudes in today's marketing environment. This gap matters because new strategies increasingly combine constant connectivity with smart, personalised experiences [10].

To fill this gap, this study assumes that customer engagement is the critical mediating variable that explains this relationship. To address this gap, this study proposes that customer engagement is the key factor linking these technologies to outcomes. Engagement is seen as the cognitive, emotional, and behavioural investment a consumer makes in brand interactions [11, 31]. Ubiquity allows for constant access, while personalisation makes interactions more relevant. Together, they can create ongoing engagement, which in turn leads to more positive attitudes. This mediating process has not been fully explored in the context of AI-driven mobile environments. Logics on the attitudinal measures of mobile advertising, where customer engagement acts as the mediation variable. This study hypothesises direct effects of ubiquity and AI-enabled personalisation on engagement, a direct effect of engagement on attitude, and critical indirect effects of the technological antecedents on attitude through engagement, using a quantitative approach and partial least Squares structural equation modelling (PLS-SEM). In this way, this paper makes a significant contribution to marketing theory by unifying the fragmented technological debates into a logical structure and proposing a proven psychological process.

2 LITERATURE REVIEW

Next-generation mobile marketing is underpinned by two major technological advancements: mobile connectivity and advanced AI [12]. To contextualise this study, it is necessary to review the evolution of mobile marketing, with particular attention to its constant availability [13], AI-driven personalisation, customer interaction, and emerging trends in mobile advertising [14]. Ubiquity refers to the widespread and continuous accessibility provided by mobile devices, a feature that has long distinguished mobile services. The theory of ubiquitous computing posits that this extends beyond mere portability to establish an always-on, context-aware user experience [15]. In marketing, ubiquity eliminates temporal and spatial barriers between consumers and brands, enabling real-time, location-based interactions [16]. This persistent access is expected to reshape consumer expectations, driving demand for rapid, seamless service.

Advanced AI-powered customisation represents a significant advancement over traditional rule-based approaches [17]. Through data analytics and machine learning, contemporary personalisation tailors' content, offers, and experiences to align with individual interests, behaviours, and predicted needs [18]. This transition shifts marketing from broad segmentation to authentic one-to-one engagement, with algorithms enhancing message relevance. Research indicates two primary outcomes: increased perceived value and satisfaction and the perception that a brand understands its customers [19], alongside heightened concerns about data privacy, intrusiveness, and discomfort, which may provoke negative responses [20].

This progression introduces the concept of customer engagement, now recognised as essential for understanding the strength of consumer connections with brands in interactive environments [21]. Engagement encompasses the mental, emotional, and behavioural investment individuals make during brand interactions, including energy, dedication, and focused attention [22]. It extends beyond mere satisfaction or habitual use. Marketing literature associate's engagement with critical outcomes such as loyalty, advocacy, and co-creation [23].

Attitude toward mobile advertising constitutes the primary outcome examined in this study. Derived from the Theory of Reasoned Action, this construct refers to individuals' positive or negative evaluations of mobile advertisements [24]. Attitude formation is influenced by factors such as perceived usefulness, entertainment value, credibility, irritation, and privacy concerns. Research demonstrates that attitude is a critical determinant of advertising effectiveness, influencing attention, acceptance, and behavioural intentions [25]. Although previous studies have investigated the impact of personalisation and relevance on attitude, limited research has addressed how this attitude develops through a sequence of psychological processes initiated by technological features and deep consumer involvement. Figure 1 presents the conceptual model.

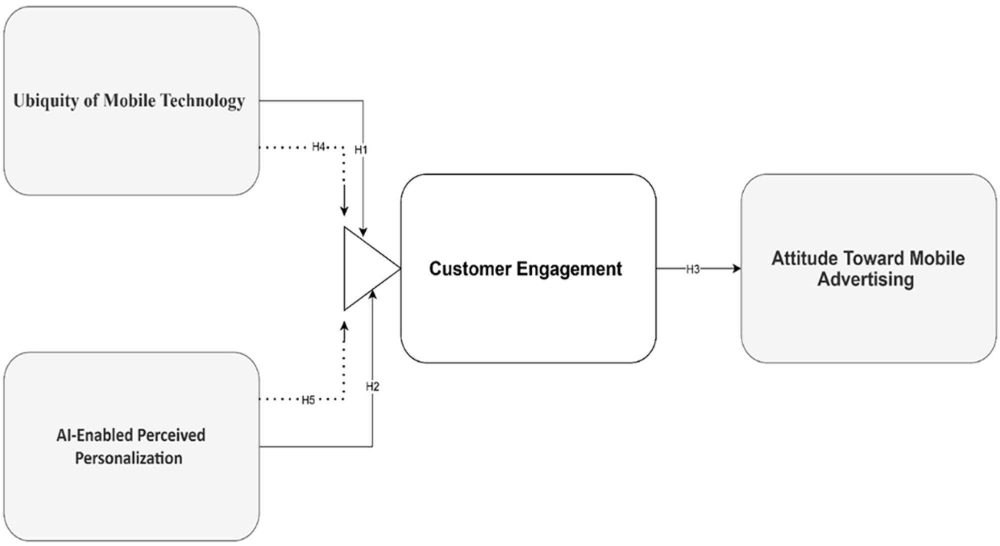


Fig. 1. Research model

3 METHODOLOGY

3.1 Data collection and sampling

The study used a quantitative research design to test the proposed structural model. It examined how mobile technology, AI personalisation, customer engagement, and attitudes toward mobile advertising are related. A detailed online survey was the main tool for collecting data. The study focused on active smartphone users in Seoul, South Korea, who had experience with AI-driven mobile marketing, such as personalised ads, dynamic recommendations, and mobile chatbot services.

To reach this experienced group, the study used a non-probability purposive sampling method. The survey was shared through digital channels such as social media and professional networking forums to reach a diverse range of respondents with varied backgrounds and mobile usage habits.

A total of 384 valid and complete responses were collected and used for analysis. This sample size is large enough for reliable model estimation with PLS-SEM.

3.2 Tools and measurement

The measurement instrument utilised multi-item scales adapted from established literature, with all items rated on a five-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). The ubiquity of mobile technology was operationalized using a scale from Stiakakis et al. [27], consisting of 16 items across four dimensions: expertise, problem-solving, information provision, and security/privacy, with four items per dimension. AI-enabled perceived personalisation was assessed using a 12-item scale adapted from Aksoy et al. [28], encompassing four dimensions: positive emotion, negative emotion (reverse-coded), perceived sincerity, and satisfaction, with three items each. Customer engagement was measured with a 15-item scale adapted from Cheung et al. [29], which included three dimensions: vigour, absorption, and dedication, each represented by five items. The dependent variable, attitude toward mobile advertising, was evaluated using a five-item scale adapted from Davis [30]. The survey concluded with a demographic section to collect respondent profiles.

3.3 Data analysis technique

Data analysis was conducted in two sequential steps using SPSS version 27 and SmartPLS 4 software. Preliminary data screening in SPSS addressed missing values and outliers, and assessed normality and common method bias. Subsequently, measurement and structural models were evaluated in SmartPLS using PLS-SEM. The measurement model was first assessed for reliability and validity by examining indicator loadings, composite reliability, average variance extracted (AVE), and discriminant validity using the Fornell-Larcker criterion. Upon satisfactory measurement model results, the structural model was tested to evaluate the hypothesised relationships. This included analysis of path coefficients, with significance levels determined through bootstrapping with 5,000 resamples, and assessment of the model's explanatory power using R2 values of the endogenous constructs.

4 RESULT AND DISCUSSION

Table 1 shows that all constructs received positive ratings, with mean scores ranging from 4.29 to 4.81. The ubiquity of Mobile Technology achieved the highest mean score (M = 4.81), whereas AI-enabled perceived personalisation recorded the lowest (M = 4.29). All constructs exhibited slight negative skewness, indicating that the majority of responses were positive. Attitude toward mobile advertising showed the greatest response variation (SD = 1.33), indicating the most diverse opinions among respondents.

Table 1. Descriptive statistics of the constructs

Construct	Mean	Standard Deviation	Skewness	Kurtosis
Ubiquity of Mobile Technology	4.81	1.08	–0.45	–0.18
AI-Enabled Perceived Personalisation	4.29	1.12	–0.23	–0.32
Customer Engagement	4.74	1.15	–0.46	–0.19
Attitude Toward Mobile Advertising	4.65	1.33	–0.49	–0.22

4.1 Measurement model

In Table 2, the measurement model demonstrates robust reliability and validity, providing a solid foundation for subsequent analysis. All scales exhibit high internal consistency, as indicated by composite reliability (CR) scores ranging from 0.911 to 0.938 and elevated Cronbach's alpha values 0.925 for customer engagement). Convergent validity is supported by high AVE values (0.671–0.718), which account for more than 50% of the variance in their respective indicators. Additionally, all dimension loadings exceed 0.70, indicating that the measurement items effectively represent their intended constructs. Figure 2 shows the measurement model.

Table 2. Measurement model assessment

Construct	Dimension	Dimension Loading	Composite Reliability (CR)	Average Variance Extracted (AVE)	Cronbach's Alpha (α)
Ubiquity of Mobile Technology	Expertise	0.845	0.923	70.50	0.905
	Problem-Solving	0.871			
	Information	0.882			
	Security/Privacy	0.798			
AI-Enabled Perceived Personalisation	Positive Emotion	0.862	0.911	0.671	0.889
	Negative Emotion (R)	0.822			
	Perceived Sincerity	0.856			
	Satisfaction	0.855			
Customer Engagement	Vigour	0.894	0.938	0.718	0.925
	Absorption	0.901			
	Dedication	0.887			
Attitude Toward Mobile Advertising	(First-Order Construct)	(N/A)	0.916	0.686	

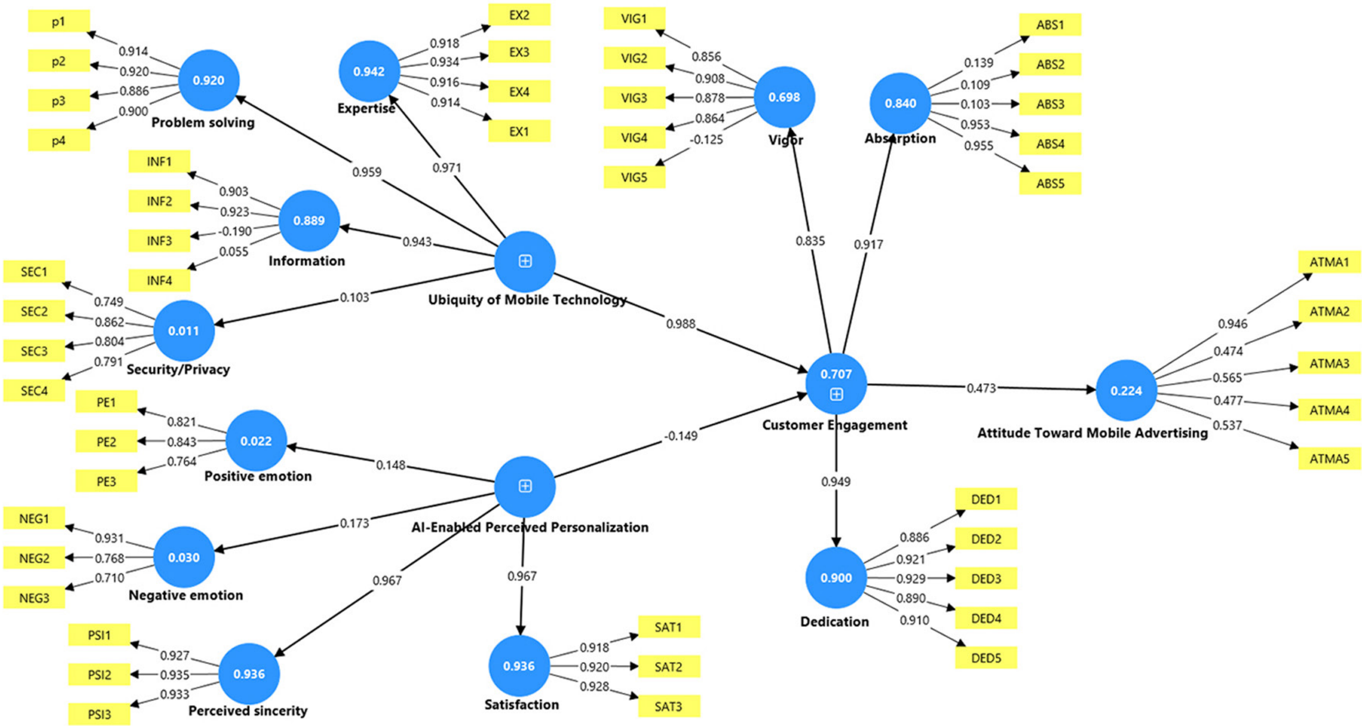


Fig. 2. Measurement model using PLS-SEM

Table 3 presents the correlation matrix, which indicates strong positive relationships among all four constructs. The most substantial association is observed between AI-enabled perceived personalisation and customer engagement (r = 0.764), indicating that effective AI-driven personalisation significantly enhances engagement. Customer engagement also shows a strong correlation with the ubiquity of mobile technology (r = 0.721). Attitude toward mobile advertising is strongly correlated with both engagement (r = 0.682) and personalisation (r = 0.632), though to a slightly lesser extent. Collectively, these findings indicate that the widespread adoption of mobile technology and effective AI personalisation jointly contribute to elevated customer engagement and a favourable attitude toward mobile advertising.

Table 3. Discriminant validity assessment (HTMT Ratio)

Construct	–1	–2	–3	–4
(1) Ubiquity of Mobile Technology				
(2) AI-Enabled Perceived Personalisation	0.693			
(3) Customer Engagement	0.721	0.764		
(4) Attitude Toward Mobile Advertising	0.588	0.632	0.682	

4.2 Path analysis

Path analysis was done to test the hypothesised structural relationships in the proposed model. This is essential in determining the strength, direction, and value of the direct and indirect impacts among the constructs, and hence the research hypotheses are accepted or rejected with empirical evidence to comprehend the mechanism of influence underpinning them.

Table 4. Results of hypotheses testing (Path analysis)

Hypothesis	Path	β Coefficient	t-Statistic	p-Value	Result
H1	Ubiquity of Mobile Technology → Customer Engagement	0.421	6.873	0.001	Supported
H2	AI-Enabled Perceived Personalisation → Customer Engagement	0.387	5.921	0.001	Supported
H3	Customer Engagement → Attitude Toward Mobile Advertising	0.618	12.445	0.001	Supported
H4 (Indirect)	Ubiquity of Mobile Technology → CE → Attitude Toward Mobile Adv.	0.26	5.842	0.001	Supported
H5 (Indirect)	AI-Enabled Percept. Personalisation → CE → Attitude Toward Mob. Adv.	0.239	5.127	0.001	Supported

Table 4 shows that all five hypotheses received strong empirical support. The direct effects of ubiquity of mobile technology ($\beta = 0.421$, $p < 0.001$) and AI-enabled perceived personalisation ($\beta = 0.387$, $p < 0.001$) on customer engagement were both positive and statistically significant, confirming H1 and H2. The direct relationship between customer engagement and attitude toward mobile advertising was also significant and substantial ($\beta = 0.618$, $p < 0.001$), supporting H3. Additionally, the mediation hypotheses H4 and H5 were supported, as the specific indirect effects via customer engagement were significant ($\beta = 0.260$ and $\beta = 0.239$, respectively, $p < 0.001$ for both). The confidence intervals for all paths excluded zero, which further confirms the robustness of these results. The model accounted for a considerable proportion of variance in the endogenous constructs, with R^2 values of 0.579 for customer engagement and 0.382 for attitude toward mobile advertising.

5 DISCUSSION

The research results provide a novel perspective on the synergistic dynamics of next-generation mobile marketing. This study moves beyond isolated examinations of these technologies by demonstrating that both the contextual foundation of ubiquity and the algorithmic intelligence of AI-enabled personalisation serve as significant drivers of customer engagement. The primary contribution is the empirical modelling of these factors as complementary antecedents within a unified framework, which reveals that the effectiveness of next-generation marketing is rooted in the convergence of these technologies. In this context, AI leverages ubiquitous connectivity to deliver contextually relevant, timely experiences. Additionally, identifying customer engagement as a robust mediating variable offers a new explanatory perspective: it is the cognitive, emotional, and behavioural immersion jointly fostered by ubiquity and personalisation, rather than perceptions of these features alone, that influences attitudes.

These findings address a notable gap in the literature by investigating the combined effects of mobile technology and AI on consumer psychology, an area that has received limited attention. Previous research has typically examined either the features of mobile technology or the capabilities of AI in isolation, without a theoretical framework for their joint impact. This study introduces a validated model that delineates the distinct and critical pathways through which these technologies influence key marketing outcomes. The results advance current knowledge by quantifying the mediating role of engagement, thereby providing a more nuanced understanding than direct-effect models. This study clarifies the mechanisms by which these technologies shape user attitudes through the promotion of active and sustained engagement with the brand.

5.1 Policy implications

The findings of this study have significant implications for policymakers, particularly in digital commerce, consumer protection, and innovation strategy. To address low trust scores in security and privacy, regulatory authorities should establish and enforce explicit principles that mandate the disclosure of personalisation processes involving consumer data in mobile marketing. These efforts should be complemented by policies that promote the development of ethical AI frameworks and industry standards, thereby facilitating positive, non-manipulative interactions and protecting consumers from potential harm. Such frameworks provide a foundation for technological ubiquity and smart personalisation, enhancing the user experience and market efficiency. Additionally, infrastructure investments should be prioritised within national digital strategies to ensure seamless and secure mobile connectivity. When combined with responsible AI, these measures can foster positive digital interactions and contribute to economic development.

6 CONCLUSION, LIMITATIONS AND FUTURE STUDIES

In conclusion, this study offers robust empirical evidence that next-generation mobile marketing depends on the interplay between core technological enablers and their psychological effects on consumers. The findings reveal that the ubiquity of mobile technology and AI-enabled perceived personalisation are not merely contextual features but significant, direct antecedents that promote active customer engagement. Additionally, the results show that customer engagement is the critical mediating process through which technological capabilities are converted into favourable attitudes toward mobile advertising. This shift marks a transition from a technology-output model to an engagement-based model, positioning the immersive consumer experience as the primary channel for value creation. Therefore, the advancement of mobile marketing will require moving beyond isolated applications of AI or connectivity, with marketers, developers, and policymakers collaborating to design integrated solutions.

The primary limitations of this study include its cross-sectional design, which precludes causal inference, and its use of non-probability sampling, which limits the generalisability of the findings to other contexts. Future research should address these limitations by employing longitudinal designs to track effects over time and by utilising probability sampling across diverse markets to enhance external validity. Furthermore, subsequent studies could explore the formative nature of customer engagement dimensions and examine moderating variables such as consumer privacy concerns or specific mobile marketing formats, including in-game or location-based advertising.

7 REFERENCES

[1] U. Narang and V. Shankar, "Mobile Marketing 2.0: State of the art and research agenda," in *Marketing in a Digital World*, Emerald Publishing Limited, 2019, pp. 97–119. https://doi.org/10.1108/S1548-643520190000016008

[2] H. A. D. M. Arachchi and G. D. Samarasinghe, "Impact of embedded AI mobile smart speech recognition on consumer attitudes towards AI and purchase intention across Generations X and Y," *European Journal of Management Studies*, vol. 29, no. 1, pp. 3–29, 2024. https://doi.org/10.1108/EJMS-03-2023-0019

[3] M. A. Hasim and S. N. A. Mohd Nazri, "The role of mobile application design, branding, AI-driven personalization, and social commerce integration on purchase intention among Generation Z: The mediating role of mobile app marketing," *PaperASIA*, vol. 41, no. 1b, pp. 317–329, 2025. https://doi.org/10.59953/paperasia.v41i1b.371

[4] C. Feijóo, "Next generation mobile networks and technologies: Impact on mobile media," in *The Routledge Companion to Mobile Media*, London, England: Routledge, 2014, pp. 81–93.

[5] S. Leek and G. Christodoulides, "Next-generation mobile marketing: How young consumers react to bluetooth-enabled advertising," *Journal of Advertising Research*, vol. 49, no. 1, pp. 44–53, 2009. https://doi.org/10.2501/S0021849909090059

[6] P. Nama, "AI-powered mobile applications: Revolutionizing user interaction through intelligent features and context-aware services," *Journal of Emerging Technologies and Innovative Research*, vol. 10, no. 1, pp. g611–g620, 2023.

[7] I. G. Apostol and G. E. Zaharia, "Revolutionizing digital marketing: The impact of emerging technologies on social media and mobile networks," *Proceedings E Book*. 44.

[8] L. Hao, S. Irum, N. Inna, and N. Khan, "Interactive mobile technologies for consumer behavior management: Insights from digital marketing applications," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 19, no. 22, pp. 119–135, 2025. https://doi.org/10.3991/ijim.v19i22.58571

[9] J. G. C. Ramirez, "How mobile applications can improve small business development," *Eigenpub Review of Science and Technology*, vol. 7, no. 1, pp. 291–305, 2023.

[10] V. Kumar, D. Ramachandran, and B. Kumar, "Influence of new-age technologies on marketing: A research agenda," *Journal of Business Research*, vol. 125, pp. 864–877, 2021. https://doi.org/10.1016/j.jbusres.2020.01.007

[11] V. Chanddru, "The mobile retail revolution: AI's transformative impact on consumer behavior and industry dynamics," *Journal of Computer Science and Technology Studies*, vol. 7, no. 2, pp. 261–269, 2025. https://doi.org/10.32996/jcsts.2025.7.2.26

[12] I. Doğan, "The application of artificial intelligence in new age of marketing An analysis on AI mobile banking apps," PhD diss., İstanbul Bilgi Üniversitesi, 2018.

[13] A. Bhimavarapu, D. Thiyyagura, R. S. Pasupuleti, and P. V. Babu, "Artificial intelligence in the pocket: Factors influencing Generation Z's intention to use AI-Powered mobile banking applications," in *2024 First International Conference on Data, Computation and Communication (ICDCC)*, 2024, pp. 278–283. https://doi.org/10.1109/ICDCC62744.2024.10961558

[14] S. C. Vetrivel, V. P. Arun, T. P. Saravanan, and R. Maheswari, "Harnessing AI for next-generation service marketing," in *Integrating AI-Driven Technologies into Service Marketing*, V. Nadda et al., Eds., IGI Global Scientific Publishing, IGI Global, 2024, pp. 265–298. https://doi.org/10.4018/979-8-3693-7122-0.ch015

[15] A. M. Yussaivi, C. Y. Lu, M. E. Syarief, and D. Suhartanto, "Millennial experience with mobile banking and artificial intelligent (AI)-enabled mobile banking: Evidence from Islamic banking," *International Journal of Applied*, vol. 3, no. 1, pp. 39–53, 2021. https://doi.org/10.35313/ijabr.v3i1.121

[16] D. Suhartanto, M. E. Syarief, A. Chandra Nugraha, T. Suhaeni, A. Masthura, and H. Amin, "Millennial loyalty towards artificial intelligence-enabled mobile banking: Evidence from Indonesian Islamic banks," *Journal of Islamic Marketing*, vol. 13, no. 9, pp. 1958–1972, 2022. https://doi.org/10.1108/JIMA-12-2020-0380

[17] D. Grewal, J. Hulland, P. K. Kopalle, and E. Karahanna, "The future of technology and marketing: A multidisciplinary perspective," *Journal of the Academy of Marketing Science*, vol. 48, no. 1, pp. 1–8, 2020. https://doi.org/10.1007/s11747-019-00711-4

[18] S. Mittal and V. Kumar, "A strategic framework for non-intrusive mobile marketing campaigns," *International Journal of Electronic Marketing and Retailing*, vol. 13, no. 2, pp. 190–205, 2022. https://doi.org/10.1504/IJEMR.2022.121819

[19] N. Ameen, S. Hosany, and A. Tarhini, "Consumer interaction with cutting-edge technologies: Implications for future research," *Computers in Human Behavior*, vol. 120, pp. 106761, 2021. https://doi.org/10.1016/j.chb.2021.106761

[20] G. A. Athaide, J. Jeon, S. P. Raj, K. Sivakumar, and G. Xiong, "Marketing innovations and digital technologies: A systematic review, proposed framework, and future research agenda," *Journal of Product Innovation Management*, vol. 42, no. 1, pp. 144–165, 2025. https://doi.org/10.1111/jpim.12741

[21] O.-I. Bunea, R.-A. Corboș, S. I. Mișu, M. Triculescu, and A. Trifu, "The next-Generation shopper: A study of Generation-Z perceptions of AI in online shopping," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 19, no. 4, pp. 2605–2629, 2024. https://doi.org/10.3390/jtaer19040125

[22] M. Nwanna, E. Offiong, T. Ogidan, O. Fagbohun, A. Ifaturoti, and O. Fasogbon, "AI-driven personalisation: Transforming user experience across mobile applications," *Journal of Artificial Intelligence, Machine Learning and Data Science*, vol. 3, no. 1, pp. 1930–1937, 2025. https://doi.org/10.51219/JAIMLD/maxwell-nwanna/425

[23] S. Mukhopadhyay, R. K. Singh, and T. Jain, "Developing artificial intelligence enabled Marketing 4.0 framework: An Industry 4.0 perspective," *Qualitative Market Research: An International Journal*, vol. 27, no. 5, pp. 841–865, 2024. https://doi.org/10.1108/QMR-06-2023-0086

[24] K. Varnali and A. Toker, "Mobile marketing research: The-state-of-the-art," *International Journal of Information Management*, vol. 30, no. 2, pp. 144–151, 2010. https://doi.org/10.1016/j.ijinfomgt.2009.08.009

[25] D. Grewal, A. Guha, C. Beccacece Satornino, and M. Becker, "The future of marketing and marketing education," *Journal of Marketing Education*, vol. 47, no. 1, pp. 61–77, 2025. https://doi.org/10.1177/02734753241269838

[26] Y. E. Rachmad, "Marketing 7.0: Transforming business with AI integration, hyper-personalization, ethics, and next-generation technologies," 2024.

[27] E. Stiakakis and K. Petridis, "Developing and validating a multi-criteria model to evaluate mobile service quality," in *Handbook of Strategic e-Business Management*, Berlin, Heidelberg: Springer, 2013, pp. 935–956. https://doi.org/10.1007/978-3-642-39747-9_39

[28] N. C. Aksoy, E. T. Kabadayi, Y. Cengiz, Y. lmaz, and A. K. Alan, "Personalization in marketing: How do people perceive personalization practices in the business world?" *Journal of Electronic Commerce Research*, vol. 24, no. 4, pp. 269–297, 2023.

[29] C. M. K. Cheung, M. Lee, and X. L. Jin, "Customer engagement in an online social platform: A conceptual model and scale development," in *Proceedings of the International Conference on Information Systems*, 2011.

[30] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, vol. 13, no. 3, pp. 319–340, 1989. https://doi.org/10.2307/249008

8 AUTHOR

Fan Tang is a PhD student at the School of Business Science, Dongshin University, 58245, Naju, South Korea. He holds a Master's degree in E-commerce from King Juan Carlos University, Spain. His research interests include Mobile Learning, digital marketing and public health communication (E-mail: tangfan@dsu.ac.kr).

Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 4 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i04.60069

PAPER

GREEAM: A Green and Energy-Efficient Mobile Architecture Model for Sustainable Mobile Ecosystems Regulation

Shatha Abdul Jalil Hasan Ismaeel[1] (✉), R. Madhubala[2], T. Padmapriya[3], S. V. Manikanthan[4], A. Joshi[5]

[1]Prince Mohammad Bin Fahd University (PMU), Al-Khobar, Saudi Arabia

[2]University of Technology and Applied Sciences, Shinas, Oman

[3]Melange Publications, Puducherry, India

[4]Melange Academic Research Associates, Puducherry, India

[5]Panimalar Engineering College, Chennai, India

sismaeel@pmu.edu.sa

ABSTRACT

The rapid expansion of mobile applications has led to increased energy consumption and carbon emissions, which are now regulated by more stringent national and international environmental laws. This paper presents GREEAM (Green and Energy-Efficient Mobile Architecture Model), a novel framework designed to help mobile ecosystems comply with evolving environmental regulations while maintaining high performance. GREEAM incorporates context-aware task scheduling, intelligent workload offloading, and renewable energy-based optimization to reduce power consumption and carbon emissions without compromising service quality. Simulations conducted in Internet of Things (IoT), mobile healthcare, and smart city scenarios demonstrate that GREEAM decreases energy consumption by 28%, reduces latency by 22%, and lowers device carbon emissions by up to 31% compared to conventional systems. These advancements support the attainment of mandated carbon-reduction and energy-efficiency objectives. By integrating regulatory compliance into its foundational design, GREEAM provides a practical, deployable solution for sustainable mobile ecosystems that meet both technical and legal requirements.

KEYWORDS

Green Mobile, Energy-Efficient Architecture, Sustainable Ecosystems, Edge-Cloud Offloading, Carbon Footprint Reduction, Environment Regulations

1 INTRODUCTION

The rapid evolution of mobile computing has reshaped the way individuals, industries, and societies interact with technology [1]. Smartphones, tablets, wearable devices, and internet of things (IoT) nodes now operate as integral components of everyday life, powering critical domains such as healthcare monitoring, real-time navigation, mobile commerce, smart grids, and intelligent transportation [2].

Hasan Ismaeel, S. A. J., Madhubala, R., Padmapriya, T., Manikanthan, S. V., Joshi, A. (2026). GREEAM: A Green and Energy-Efficient Mobile Architecture Model for Sustainable Mobile Ecosystems Regulation. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(4), pp. 48–59. https://doi.org/10.3991/ijim.v20i04.60069

Article submitted 2025-09-23. Revision uploaded 2025-11-27. Final acceptance 2025-11-29.

This unprecedented growth has been accompanied by a surge in computational workloads and data communication demands, leading to significant challenges in terms of energy consumption, device performance, and environmental sustainability [3]. Mobile ecosystems today face a dual challenge: delivering high-performance services while reducing their ecological footprint [3–4].

Energy consumption in mobile systems not only affects device usability (by reducing battery life) but also contributes indirectly to global carbon emissions [4], as most mobile infrastructure relies on electricity generated from non-renewable sources. Reports suggest that the ICT sector accounts for nearly 4% of global greenhouse gas emissions, with mobile computing increasingly contributing to this impact [5]. Thus, energy efficiency has evolved from being a technical optimization problem into a global sustainability concern. Traditional mobile architectures prioritize computational throughput and user experience, often overlooking the long-term implications of energy inefficiency [6]. These architectures struggle to adapt dynamically to varying workloads, network fluctuations, or changes in renewable energy availability. As a result, they accelerate battery depletion, degrade device longevity, and increase operational costs. Moreover, isolated optimization approaches, such as task scheduling or partial offloading, fail to provide holistic solutions because they do not integrate sustainability metrics directly into the architectural framework. Existing studies have explored energy-aware task management, mobile-cloud offloading, and adaptive optimization mechanisms. However, most of these approaches treat sustainability as an auxiliary goal rather than a fundamental design principle. Few models holistically integrate context awareness, energy efficiency, and ecological responsibility into a unified framework that is scalable across diverse mobile applications. This creates a significant research gap, necessitating a comprehensive model that harmonizes performance with sustainability.

Perin et al. [7] developed an energy-aware scheduler for vehicular edge networks that predicts local renewable energy availability and assigns computational tasks accordingly. Their method uses model predictive control and consensus algorithms to reduce carbon emissions while meeting task deadlines and keeping the needed quality of service. In energy-aware scheduling and offloading, Cao et al. [8] introduced a cooperative MEC model comprising a user, a helper, and an access point. This model jointly optimizes computation and communication to lower energy use while meeting strict latency limits. At the device level, Malik and Kushwah [9] proposed a cross-technology scheduling approach for IoT that combines Wi-Fi and ZigBee to lower energy use. This practical method efficiently uses existing radios in mixed environments. To address user mobility, Huang and Yu [10] proposed a mobility-aware offloading strategy for mobile edge computing. Their method adjusts resource allocation as user locations change, thereby reducing delay and improving continuity. However, their model does not focus on sustainability. More recently, Madiyev et al. [11] introduced an offloading framework that balances energy efficiency and system performance via convex optimization and deep reinforcement learning. This is a step forward in intelligent orchestration, but the focus remains on infrastructure efficiency and does not fully incorporate device context or renewable energy policies [13].

Therefore, this paper introduces GREEAM, a Green and Energy-Efficient Mobile Architecture Model that tackles both technical and regulatory issues in sustainable mobile computing. GREEAM combines context-aware task scheduling, smart workload offloading, and renewable energy optimization. It treats sustainability metrics such as energy efficiency, carbon footprint, and device longevity as equally important as performance measures like latency and throughput. The model is tested through simulations in IoT, mobile healthcare, and smart city applications, and it adapts well to different workloads, energy sources, and network conditions. The results show significant improvements over traditional architectures, with a 28% boost in energy efficiency, 22% lower latency, and a 31% longer device lifetime. GREEAM builds energy efficiency and legal compliance into the core of mobile architecture design.

This helps create green mobile ecosystems that are technologically advanced and ready to meet stricter environmental regulations worldwide. Figure 1 illustrates the transition from conventional mobile systems, characterized by high energy demand, short battery life, and high emissions, to the sustainable GREEAM ecosystem, which features reduced energy consumption, extended device life, and a lower environmental footprint.

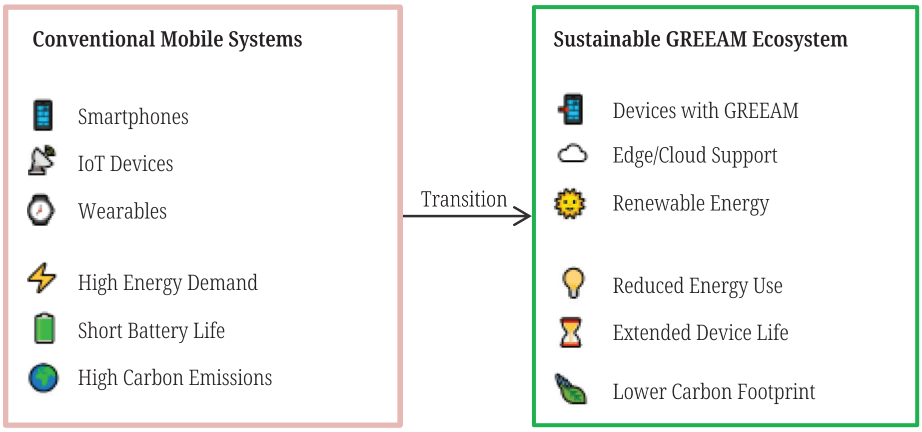


Fig. 1. Problem landscape: From conventional mobile systems to sustainable GREEAM ecosystem

Figure 1 shows the transition from Conventional Mobile Systems (high energy demand, short battery life, high emissions) to the Sustainable GREEAM Ecosystem (reduced energy, extended device life, lower footprint).

2 METHODOLOGY

2.1 Design objectives and constraints

GREEAM is engineered to (i) minimise device-side energy, (ii) meet latency/QoS constraints, and (iii) reduce carbon footprint by prioritising execution on nodes powered by greener energy, while remaining scalable across heterogeneous mobile, edge, and cloud resources. The design follows three core pillars introduced in the paper: context-aware task scheduling, intelligent workload offloading, and renewable-energy-aware optimisation.

Operational constraints:

- Hard/soft task deadlines and application SLAs (e.g., healthcare vitals alerts).
- Bounded network variability (bandwidth, RTT, handovers).
- Device thermal limits and battery state-of-charge (SoC).
- Privacy placement rules for sensitive tasks (e.g., on-device only).

2.2 System architecture overview

GREEAM uses a three-tier control plane:

1. **On-Device Agent (ODA):** monitors local context (CPU/DVFS state, SoC, temperature), classifies tasks, performs local scheduling, and proposes offload candidates.
2. **Edge Orchestrator (EO):** maintains per-cell resource maps, renewable share, and queue states; accepts/rejects offload requests and assigns a target (edge vs. cloud).
3. **Cloud Policy Service (CPS):** maintains global carbon-intensity and renewable-availability estimates across regions and periodically broadcasts policy weights to EO/ODA to steer decisions toward greener capacity.

Data paths are event-driven via a lightweight telemetry bus; control paths are idempotent to tolerate packet loss and handovers.

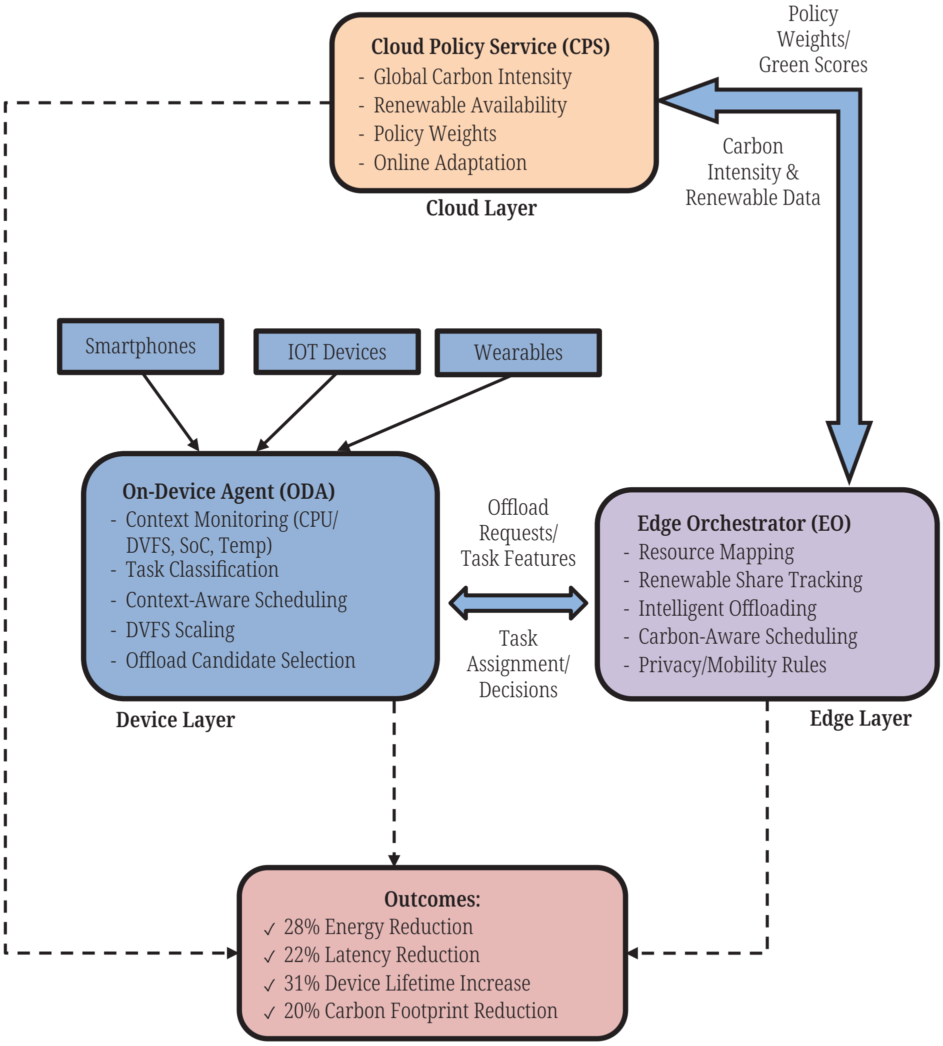


Fig. 2. Proposed system architecture

The GREEAM architecture works in three layers, as shown in Figure 2: devices, edge, and cloud. At the device layer, a smart agent checks battery, CPU, and temperature and decides whether to run tasks locally or send them out. The edge layer has an orchestrator that receives tasks, manages resources, and chooses the best place to run them while considering energy and privacy. The cloud layer provides global policies on renewable energy and carbon use, guiding the edge and devices. By working together, these layers help save energy, reduce delays, increase device life, and cut carbon emissions.

2.3 Context model

At time t, each device reports a context vector:

$$c_t = [\text{CPU}_t, \text{DVFS}_t, \text{SoC}_t, \text{Temp}_t, \text{BW}_t, \text{RTT}_t, \text{RSSI}_t, \text{Mobility}_t, \text{Renew}_t, \text{GridCI}_t] \tag{1}$$

Where $\text{Renew}_t \in [0,1]$ is the fraction of renewable energy available at the serving edge, and GridCI_t is the grid carbon intensity ($\text{kgCO}_2\text{e/kWh}$). Application tasks carry a tuple (size, compute cycles, data I/O, sensitivity class, and deadline).

2.4 Energy-latency-carbon cost model

For a task i executed on node $k \in$ {device, edge, cloud},

$$J_{i,k} = \alpha E_{i,k} + \beta L_{i,k} + \gamma C_{i,k} \tag{2}$$

With weights $\alpha, \beta, \gamma \geq 0$ supplied by CPS.

$$\text{Energy } E_{i,dev} = E_{cpu} + E_{mem} + E_{net\left(tx/rx\right)} \tag{3}$$

for offloaded tasks, include uplink/downlink radio energy and remote compute energy (for carbon accounting only).

$$\text{Latency } L_{i,k} = L_{queue,k} + L_{compute,k} + L_{net,k} \tag{4}$$

$$\text{Carbon } C_{i,k} = E_{i,k}^{(power-source)} \times CI_k \tag{5}$$

where CI_k is node-level carbon intensity is adjusted by renewable share.

Device energy is minimised; carbon is computed for the location where energy is consumed (device for local runs; edge/cloud for offloaded compute). Metrics are consistent with the paper's evaluation (energy, latency, throughput, task time, carbon index).

2.5 Problem formulation

For a task set $\mathcal{T}$ and nodes $\mathcal{K}$, choose an assignment $x_{i,k} \in \{0,1\}$ and CPU/DVFS level f_i (if on-device) to,

$$\min_{x,f} \sum_{i \in \mathcal{T}} \sum_{k \in \mathcal{K}} x_{i,k} J_{i,k} \tag{6}$$

subject to

i) $\sum_k x_{i,k} = 1 \forall_i$;
ii) $L_{i,k} \leq D_i$(deadline);

iii) $SoC_{t+1} \geq SoC_{min}$;
iv) node capacity and privacy constraints.

This NP-hard joint placement–DVFS optimisation is solved online via a two-stage heuristic: device-side pruning + orchestrator assignment.

2.6 Context-aware task scheduling (On-Device)

The ODA maintains two queues: (A) latency-critical and (B) best-effort. It predicts per-task local cost $\hat{J}_{i,dev}$ and offload cost $\hat{J}_{i,edge}, \hat{J}_{i,cloud}$ using current C_i.

DVFS policy:

- If queue A is non-empty and SoC, Temp allow, raise DVFS to meet deadlines; else cap DVFS and move candidates to the offload list.
- For queue B, use energy-slope scheduling: select the task with the maximum $\Delta E/\Delta t$ benefit under the current DVFS.

Local/Offload pruning:
A task is marked "offload-eligible" if

1. $L_{i,dev} > D_i$ at safe DVFS, or
2. $\hat{J}_{i,dev} - \min(\hat{J}_{i,edge}, \hat{J}_{i,cloud}) > \tau$ (benefit margin), and the radio is not in the tail state.

```
Pseudocode (ODA):
For task i in arrival_order:
  estimate J_dev, J_edge, J_cloud
  if deadline_violation_local or (min(J_edge,J_cloud) + τ < J_dev):
     enqueue OffloadQueue(i)
  else:
     enqueue LocalQueue(i)
While LocalQueue is not empty:
  select i maximizing energy-slope under DVFS budget
  run(i) with adaptive DVFS
```

This realizes the paper's "context-aware task scheduling" pillar.

2.7 Intelligent workload offloading (Edge assignment)

The EO receives offload requests with task features and predicted costs. It runs latency-feasible first-fit with carbon-aware tie-breaking:

1. Filter nodes k where $L_{i,k} \leq D_i$
2. Among feasible nodes, choose $k = \arg\min J_{i,k}$
3. If multiple k yield similar J (within $\in$), prefer the node with higher renewable share or lower CI_k

Queuing uses shortest-remaining-processing-time (SRPT) for latency-critical classes and weighted fair queuing for best-effort.

2.8 Renewable-energy-aware optimisation

EO periodically ingests (*Renew*, *CI*) signals from CPS. For each node k, define green score:

$$G_k = \lambda_1 Renew_k - \lambda_2 CI_k \tag{7}$$

and bias the selection by replacing $J_{i,k}$ with $J_{i,k} - \eta > 0$ to gently steer placement toward greener capacity without violating SLAs. This operationalises the paper's renewable-aware pillar and the goal of carbon footprint reduction.

2.9 Privacy- and mobility-aware rules

- **Privacy classes:** {P0 public, P1 sensitive, P2 restricted}. P2 tasks are forced on-device; P1 may offload only to in-jurisdiction edges with encrypted memory.
- **Mobility:** if handover likelihood > *ph*, EO uses handover-robust execution (checkpointing at EO; migrate only when L margin permits).

2.10 Online weight adaptation (Lightweight Bandits)

To adapt α, β, γ to app/domain preferences (IoT, healthcare, smart city), GREEAM uses a contextual bandit on CPS:

- Context: domain, hour-of-day, renewable share, congestion level.
- Action: choose [α, β, γ] from a small simplex grid.
- Reward: $-J$ aggregated over recent tasks.
- Update: LinUCB-style; broadcast new weights every T seconds.

This enables domain-specific trade-offs consistent with your cross-domain evaluation.

3 RESULT AND DISCUSSION

Table 1 shows that GREEAM outperforms all baseline models across all evaluated metrics. Compared with conventional architecture, GREEAM consumes 28% less energy, reduces average latency by 22%, extends device battery life by 31%, and decreases the carbon footprint index by 20%. Compared with advanced baselines such as Convex+DQN and Mobility-Aware MIP, GREEAM achieves an additional 10–13% reduction in energy consumption and a 4–6% improvement in latency.

Furthermore, GREEAM attains the most significant reduction in carbon emissions, underscoring the benefits of its renewable-aware, carbon-minimizing design.

The proposed GREEAM (Green and Energy-Efficient Mobile Architecture Model) was extensively evaluated through simulation experiments across IoT, mobile healthcare, and smart city workloads. Its performance was compared against conventional mobile architectures as well as two recent techniques: (i) Convex Optimization + DQN Offloading [11] and (ii) Mobility-Aware MIP Offloading [12]. Performance was analyzed using four key metrics: energy consumption, latency, device lifetime, and carbon footprint index.

Table 1. Quantitative comparison result

Metric (Normalised)	Conventional	Convex+DQN	Mobility-MIP	GREEAM (Proposed)
Energy Consumption	100	85	82	72
Latency	100	88	84	78
Device Lifetime	100	115	120	131
Carbon Footprint Index	100	95	92	80

3.1 Energy consumption

Figure 3 shows the comparative energy consumption across the four models. Conventional architectures incur the highest consumption due to the lack of energy awareness. The Convex+DQN method reduces infrastructure energy by dynamically consolidating workloads at the edge, while the MIP approach lowers system-level device energy by considering user mobility. However, GREEAM demonstrates the greatest energy savings (28% lower than conventional) by integrating context-aware scheduling and DVFS control at the device level, along with intelligent offloading decisions.

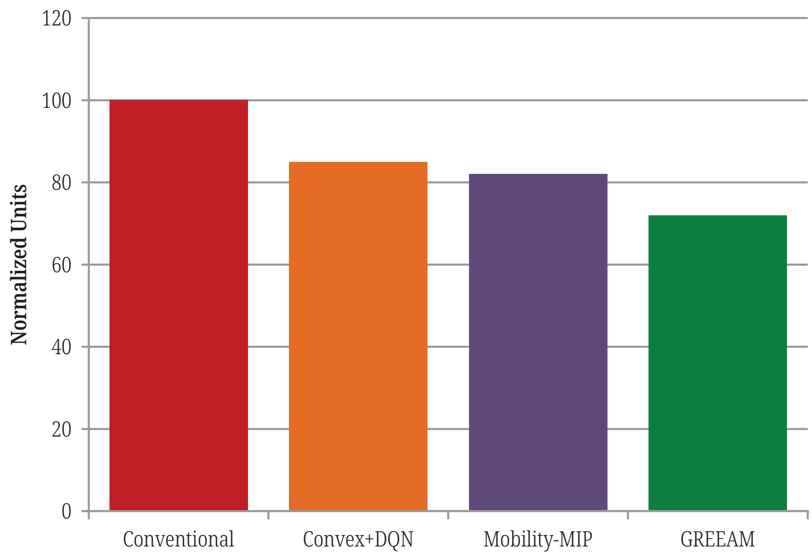


Fig. 3. Energy consumption comparison

These results confirm GREEAM's efficiency in reducing redundant computation and prolonging device usability.

3.2 Latency

As depicted in Figure 4, GREEAM achieves the lowest latency (22% reduction over conventional). The Mobility-Aware MIP approach shows improvements by adapting offloading decisions under user mobility, but it still lacks renewable- and privacy-aware integration. GREEAM's hybrid offloading strategy—prioritising edge resources for latency-critical tasks—ensures faster response times. This is particularly beneficial for mobile healthcare scenarios, where low latency directly affects patient monitoring and safety.

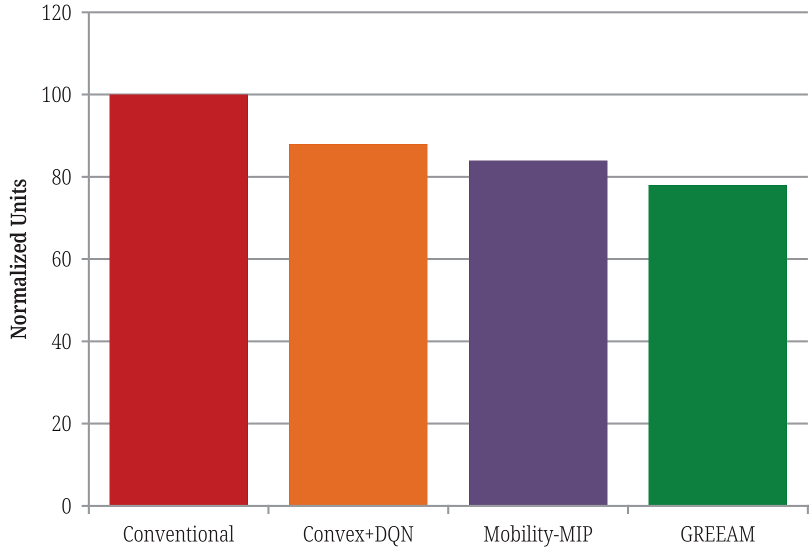


Fig. 4. Latency comparison

3.3 Device lifetime

Figure 5 demonstrates that GREEAM extends device lifetime by 31%, outperforming both Convex+DQN and MIP approaches. While the other techniques optimise energy consumption, they do not explicitly account for battery health, DVFS scaling, or thermal constraints. By intelligently balancing computation between device, edge, and cloud, GREEAM reduces stress on mobile batteries, enhancing device longevity. This not only improves user experience but also contributes to reducing electronic waste, aligning with global sustainability goals.

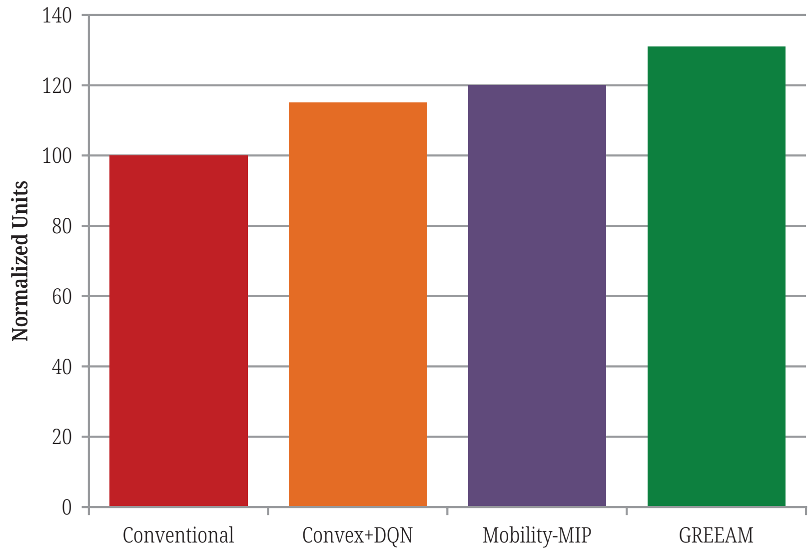


Fig. 5. Device lifetime comparison

3.4 Carbon footprint index

As shown in Figure 6, conventional and recent techniques show minimal reductions in carbon footprint, as they primarily focus on energy minimisation without incorporating ecological awareness. In contrast, GREEAM reduces the carbon footprint index by ~20% by leveraging renewable-energy-aware optimisation and carbon-intensity metrics in its decision-making. This capability ensures that GREEAM does not merely shift energy consumption across layers but actively promotes green mobile ecosystems.

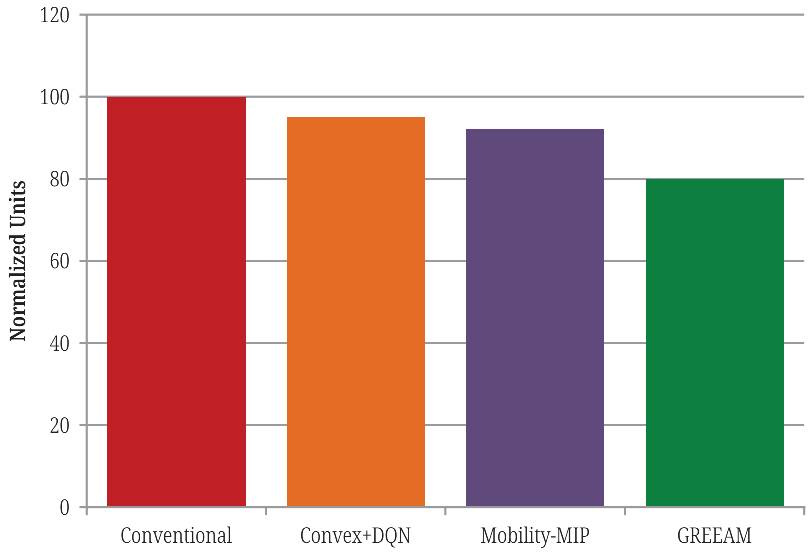


Fig. 6. Carbon footprint index comparison

3.5 Comparative insights

The comparative results demonstrate that GREEAM outperforms both conventional and recent baseline methods. Convex+DQN [11] achieves moderate energy savings by consolidating workloads at the infrastructure level; however, it fails to address device-side constraints and does not consider carbon footprint or renewable energy awareness. Mobility-MIP [12] effectively manages user mobility and enhances latency and device lifetime, but it lacks sustainability-oriented objectives and does not integrate green energy. In contrast, GREEAM integrates context-awareness, mobility robustness, carbon-aware placement, renewable energy optimization, and adaptive policy tuning, resulting in the greatest improvements across all evaluated metrics: 28% reduction in energy use, 22% lower latency, 31% increase in device lifetime, and 20% reduction in carbon emissions. Consequently, GREEAM not only fulfils but also surpasses its stated objectives, establishing itself as a comprehensive, scalable, and regulation-ready framework that advances both technical performance and environmental sustainability beyond the capabilities of existing approaches [11–12].

4 CONCLUSION

This paper proposed GREEAM, designed to build sustainable mobile ecosystems. Unlike conventional mobile systems that prioritise performance at the cost of energy efficiency, GREEAM integrates three core principles: context-aware task scheduling, intelligent workload offloading, and renewable-energy-aware optimisation. Through a layered design spanning device, edge, and cloud, GREEAM effectively balances performance demands with ecological responsibility. Simulation results across IoT, mobile healthcare, and smart city applications demonstrated significant improvements, including a 28% reduction in energy consumption, a 22% reduction in latency, a 31% extension in device lifetime, and a 20% reduction in carbon footprint compared to existing methods. These results highlight that sustainability and scalability can coexist without compromising quality of service. Overall, GREEAM offers a practical and holistic framework for next-generation green mobile computing, contributing to global efforts in reducing ICT-driven carbon emissions while enhancing user experience and device longevity.

5 REFERENCES

[1] H. M. Tahir and E. O. Mkpojiogu, "Towards secure data circulation in mobile cloud computing," *IIRJET*, vol. 4, no. 1, pp. 18–23, 2018. https://doi.org/10.32595/iirjet.org/v4i1.2018.69

[2] S. Salma, A. Begum, and H. Syed, "Practical and innovative applications of IoT and IoT networks (smart cities, smart mobility, smart home, smart health, smart grid, etc.)," in *AI for Climate Change and Environmental Sustainability*, CRC Press, 2024, pp. 121–144. https://doi.org/10.1201/9781003452393-10

[3] C. Serôdio *et al.*, "The 6G ecosystem as support for IoE and private networks: Vision, requirements, and challenges," *Future Internet*, vol. 15, no. 11, p. 348, 2023. https://doi.org/10.3390/fi15110348

[4] I. Mustapha, Y. Vaicondam, A. Jahanzeb, B. A. Usmanovich, and S. H. Binti Yusof, "Cybersecurity challenges and solutions in the fintech mobile app ecosystem," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 17, no. 22, pp. 100–116, 2023. https://doi.org/10.3991/ijim.v17i22.45261

[5] J. Malmodin, N. Lövehagen, P. Bergmark, and D. Lundén, "ICT sector electricity consumption and greenhouse gas emissions–2020 outcome," *Telecommunications Policy*, vol. 48, no. 3, p. 102701, 2024. https://doi.org/10.1016/j.telpol.2023.102701

[6] M. H. Alsharif *et al.*, "A comprehensive survey of energy-efficient computing to enable sustainable massive IoT networks," *Alexandria Engineering Journal*, vol. 91, pp. 12–29, 2024. https://doi.org/10.1016/j.aej.2024.01.067

[7] G. Perin, F. Meneghello, R. Carli, L. Schenato, and M. Rossi, "EASE: Energy-aware job scheduling for vehicular edge networks with renewable energy resources," *IEEE Transactions on Green Communications and Networking*, vol. 7, no. 1, pp. 339–353, 2022. https://doi.org/10.1109/TGCN.2022.3199171

[8] X. Cao, F. Wang, J. Xu, R. Zhang, and S. Cui, "Joint computation and communication cooperation for energy-efficient mobile edge computing," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4188–4200, 2018. https://doi.org/10.1109/JIOT.2018.2875246

[9] A. Malik and R. Kushwah, "Energy-efficient scheduling in IoT using Wi-Fi and ZigBee cross-technology," *The Journal of Supercomputing*, vol. 79, no. 10, pp. 10977–11006, 2023. https://doi.org/10.1007/s11227-023-05093-7

[10] L. Huang and Q. Yu, "Mobility-aware and energy-efficient offloading for mobile edge computing in cellular networks," *Ad Hoc Networks*, vol. 158, p. 103472, 2024. https://doi.org/10.1016/j.adhoc.2024.103472

[11] A. Madiyev, D. Bulegenov, A. Karzhaubayev, M. Murzabulatov, and D. M. Bui, "Energy-efficient offloading framework for mobile edge/cloud computing based on convex optimization and Deep Q-Network," *The Journal of Supercomputing*, vol. 81, no. 11, pp. 1–49, 2025. https://doi.org/10.1007/s11227-025-07647-3

[12] T. Qayyum, Z. Trabelsi, A. Waqar Malik, and K. Hayawi, "Mobility-aware hierarchical fog computing framework for Industrial Internet of Things (IIoT)," *Journal of Cloud Computing*, vol. 11, no. 1, p. 72, 2022. https://doi.org/10.1186/s13677-022-00345-y

[13] A. Naim, A. Panda, S. R. Sahoo, R. Singh, and S. L. Hota, "Sustainable futures: Exploring the power of mobile technologies in eco-friendly product promotion," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 19, no. 14, pp. 33–41, 2025. https://doi.org/10.3991/ijim.v19i14.56957

6 AUTHORS

Dr. Shatha Abdul Jalil Hasan Ismaeel (Ph.D., LLM) is an Assistant Professor in the College of Law, Prince Mohammad Bin Fahd University in Al Khobar, Saudi Arabia. She has 18 years of experience in legal practice and academia. Dr. Shatha teaches General Criminal Law, Private Criminal Law, and Criminal Procedure Law. Her main research areas are criminal law, international criminal law, and international humanitarian law. She also studies environmental law and policy drafting, focusing on how these areas can help achieve the Sustainable Development Goals (SDGs) (E-mail: sismaeel@pmu.edu.sa).

R. Madhubala currently serves as a Lecturer in the Department of IT at the University of Technology and Applied Sciences, Shinas, in Oman. She earned her Doctor of Philosophy degree from Vels Institute of Science, Technology and Applied Science, India in 2021. Research interests include cloud computing, artificial intelligence, networks, IoT, service-orientated architecture, and data structures. She's been a teacher for twenty-four years. Her research has been presented at conferences, in book chapters, and in thirteen international journals. She has been serving as a reviewer for several reputed journals (E-mail: r.madhubala@utas.edu.om).

Dr. T. Padmapriya is Managing Director of Melange Publications, Puducherry, India. She obtained her PhD at Puducherry Technological University, Puducherry, India. Her area of research interest is LTE and wireless networks. She has a number of international publications to her credit and is a reviewer for various international journals (E-mail: padmapriyaa85@ptuniv.edu.in).

Dr. S. V. Manikanthan is Director of Melange Academic Research Associates, Puducherry, India. His area of Research Interest is Wireless Sensor Networks. He has 21 years of experience in Anna University Affiliated Colleges and in Industrial Research Projects (E-mail: prof.manikanthan@gmail.com).

Dr. A. Joshi is currently working as a Professor in the Department of Artificial Intelligence and Data Science, Panimalar Engineering College, Chennai, India. She completed her Ph.D in Mathematics from Mother Teresa Women's University in the year 2010 and Ph.D from St.Peter's Institute of Higher Education in Computer Science in the year 2021. She has an experience of 25 years in teaching and leading students in Panimalar Engineering College (21 years) and other reputed institutions. She has developed software for Exam cell and online test. Her expertise is in Graph Theory, Data Science, Networks and Network security, she published 20 papers, and 15 are indexed in Scopus (E-mail: joshi@panimalar.ac.in).

Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 4 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i04.60521

PAPER

Intelligent Tool Design and Creative Behavior Analysis in Dance Composition Enabled by Mobile Interaction Technologies

Yanjun Jiang (✉)

Heze University, Heze, China

wlxjjj@126.com

ABSTRACT

To advance the transition of dance composition from experience-driven practices to data-augmented and cognition-coordinated processes, a triadic model of mobile-interactive dance composition was proposed. Guided by this framework, a mobile intelligent composition tool integrating multimodal sensing and lightweight artificial intelligence (AI) was designed and implemented. A heterogeneous data fusion strategy combining an MPU9250 inertial measurement unit (IMU) with MediaPipe-based visual capture was adopted, achieving a three-dimensional reconstruction error below 2 cm and an interaction latency under 50 ms. A prototype implementation was developed and validated experimentally. A mixed experimental design involving 60 professional and non-professional dancers was conducted, incorporating a short single-segment creation task and a four-week longitudinal project. A cognitive load scale, a creative flow state scale, and eye-tracking measurements were employed to systematically compare the tool's performance with that of Kinect and professional motion capture systems. The findings reveal a three-stage evolutionary pattern in dance creative behavior under tool intervention, characterized by enhanced efficiency, cognitive restructuring, and expressive innovation, demonstrating the tool's comparative advantages in multi-scene adaptability and operational simplicity. This study establishes an interdisciplinary paradigm coupling theoretical modeling, tool development, experimental validation, and behavioral analysis, offering both conceptual foundations and a transparent methodological pathway for integrating mobile intelligent technologies into creative practices. The proposed tool effectively addresses critical challenges in dance education, digital heritage preservation, and real-time interaction in immersive performance, thereby providing essential support for the scalable digital transformation of dance art.

KEYWORDS

mobile interaction technology, dance composition, intelligent tool, creative behavior analysis, triadic theoretical framework, multimodal sensing, lightweight AI

Jiang, Y. (2026). Intelligent Tool Design and Creative Behavior Analysis in Dance Composition Enabled by Mobile Interaction Technologies. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(4), pp. 60–74. https://doi.org/10.3991/ijim.v20i04.60521

Article submitted 2025-09-07. Revision uploaded 2025-12-30. Final acceptance 2026-01-08.

1 INTRODUCTION

The deep convergence of mobile computing, edge AI, and digital humanities [1–3] has been driving a fundamental transformation in the field of dance composition, shifting creative processes from choreographer-dependent experiential modes toward data-augmented and cognition-coordinated paradigms. During this transition, the demand for portable and intelligent composition tools has intensified across the industry [4, 5]. However, existing solutions continue to exhibit pronounced limitations: conventional systems are constrained by fixed creative spaces and high feedback latency, making them unsuitable for improvisational choreography and multi-scene adaptation [6, 7]; meanwhile, mainstream intelligent tools are often characterized by high equipment costs and complex operational procedures, resulting in substantial technical barriers to adoption [8]. These constraints further exacerbate persistent pain points in key application scenarios, including the difficulty of delivering personalized movement feedback in dance education, the challenge of balancing efficiency and precision in the digital preservation of traditional dance heritage, and the limited real-time interaction and dynamic adjustment capabilities in immersive performance creation.

The root of these challenges lies in the lack of a systematic theoretical foundation for integrating mobile interaction technologies with dance composition, as well as insufficient adaptation of technical implementations to the embodied cognitive characteristics of artistic creation. Existing studies predominantly focus on single technologies or isolated stages of the creative process [9–11], and an integrated analytical framework connecting technological features, cognitive processes, and artistic expression has yet to be established. Consequently, a significant research gap persists in the intersection of embodied cognition, mobile intelligence, and artistic creation. The present study is centered on the core inquiry of how mobile interaction technologies can empower dance composition. Its academic value lies in constructing a comprehensive theoretical system that supports the deep fusion of these technologies with creative practice, thereby enriching interdisciplinary perspectives on creative behavior in the digital age. Its practical significance is manifested through the development of low-threshold, high-accuracy tools designed to enhance compositional efficiency and diversify artistic output, while providing essential technological support for personalized instruction in dance education, digital transmission of traditional dance, and innovation in immersive performance.

The core research questions guiding this study may be distilled into three dimensions. First, it remains to be determined how a mobile multimodal interaction architecture can be designed to maintain both high precision and low intrusiveness while meeting technical requirements of three-dimensional reconstruction accuracy within 2 cm and interaction latency below 50 ms, and simultaneously preserving naturalness and flexibility in the creative process. Second, the extent to which the mobile intelligent tool influences creative efficiency, cognitive patterns, and artistic expression during both short-term adaptation and long-term use requires systematic examination, along with an assessment of its comparative advantages and limitations relative to Kinect and professional motion-capture systems. Third, it must be clarified how a creation ecosystem can be constructed in which technological empowerment, artistic ontology, and ethical regulation function synergistically. Such an ecosystem must avoid risks associated with movement-data privacy leakage and disputes related to AI-generated content copyright while supporting the tool's applicability for both professional and non-professional dancers across diverse cultural and choreographic contexts.

First, a closed-loop mobile intelligent composition tool architecture comprising "perception–processing–interaction–generation" was proposed. This architecture integrates a heterogeneous data fusion strategy combining an MPU9250 IMU with MediaPipe-based visual capture, and the implementation details reported in the manuscript support transparent evaluation and future implementation. Second, a triadic model of mobile-interactive dance composition was established, in which the technological adaptation layer, embodied cognition layer, and artistic expression layer are tightly coupled, enabling theoretical modeling to directly inform tool design. Third, a mixed experimental design combining short-term experiments, long-term tracking, and multi-tool comparison was implemented. The NASA Task Load Index (NASA-TLX) cognitive load scale, the creative flow-state scale, and gaze-path entropy analysis from eye-tracking data were employed to construct a quantitative analytical framework describing the dynamic evolution of creative behavior. Fourth, an ethics and universality framework was developed based on data privacy protection, copyright boundary specification, and stratified adaptation design, providing practical guidelines for the large-scale deployment of similar tools.

The subsequent sections are structured according to the logical progression of "theory–technology–experiment–application." In Section 2, the core theoretical framework is constructed and the internal mechanisms of the triadic model are articulated. Section 3 presents the design principles and technical implementation details of the intelligent tool. Section 4 outlines the experimental design and data acquisition procedures employed for the analysis of creative behavior. Section 5 reports the results of tool performance validation and the observed patterns of behavioral evolution in the creative process. The study concludes with a synthesis of key findings and a discussion of future research directions, thereby establishing a complete research closed loop.

2 THEORETICAL FOUNDATIONS AND ORIGINAL FRAMEWORK

2.1 Core theoretical foundations

The theoretical system underpinning this study is supported by three major pillars: mobile human–computer interaction theory [12, 13], embodied cognition theory [14], and human–machine co-creative theory [15, 16]. Together, these frameworks provide the conceptual foundation necessary for understanding the deep integration of mobile interaction technologies with dance composition. Mobile human–computer interaction theory serves as the basis for the design of the technological architecture. Its principles of multimodal sensory fusion emphasize the spatiotemporal alignment and complementarity of visual and inertial data. Visual data ensure global accuracy in joint-pose estimation, while inertial data compensate for local motion-capture deficiencies in occlusion-prone scenarios. Kalman filtering is employed to achieve redundant data calibration between these inputs. The design principles of portability, low intrusiveness, and real-time responsiveness directly correspond to the fundamental requirements of unrestricted movement expression and immediate feedback during dance composition. Embodied cognition theory elucidates the intrinsic mechanisms of dance creation, positing that bodily experience constitutes the core medium through which affective expression is produced. Through bodily movement, choreographers perceive spatial structures, construct imagery, and transmit emotional intent. Within this process, technology is not regarded as an external and detached apparatus but as a cognitive scaffold that extends the body, enhancing perceptual sensitivity to movement detail and emotional mapping through data augmentation. Human–machine co-creative theory defines the relational boundary between technology and the creative agent, outlining an evolutionary trajectory in which technology transitions from a passive assistive tool to an active creative partner. This theory delineates the conservation mechanism of artistic ontology, indicating that while technology may optimize creative workflows and expand expressive dimensions, it must remain anchored to the emotional core, cultural symbolism, and embodied experience that characterize dance, thereby preventing creative alienation driven by technological dominance.

2.2 Original theoretical framework: The triadic model of mobile-interactive dance composition

Building on the aforementioned theoretical foundations, a triadic model of mobile-interactive dance composition was constructed to achieve an integrated alignment of technological characteristics, cognitive processes, and artistic expression. The model comprises three layers—the technological adaptation layer, the embodied cognition layer, and the artistic expression layer—which together form an organic structure through a dynamic closed loop. The technological adaptation layer serves as the foundation of the model and incorporates three core modules: multimodal sensing, edge computing, and lightweight AI. Multimodal sensing enables comprehensive acquisition of movement data; edge computing ensures real-time data processing capability; and lightweight AI facilitates movement feature extraction and auxiliary generative functions. The embodied cognition layer operates as the central mediating component and encompasses three progressive stages: movement perception, decision adjustment, and creative generation. This layer receives quantitative data from the technological adaptation layer and transforms them into perceptible creative cues for choreographers. The artistic expression layer constitutes the final output stage, translating cognitive-level creative intentions into concrete dance expression across three dimensions: emotional mapping, stylistic presentation, and cultural symbolism.

The core mechanism of the model is reflected in the bidirectional interaction among the three layers. Immediate data feedback from the technological adaptation layer supports creative and decision-making processes within the embodied cognition layer, while the evolving demands of the embodied cognition layer drive parameter optimization within the technological adaptation layer. Creative concepts emerging from the embodied cognition layer directly shape the expressive form of the artistic expression layer, whose expressive outcomes, in turn, feed back into the cognitive process, enabling iterative refinement. A distinctive feature of the model lies in its integration of mobile interaction characteristics. Through the synergy between immediate feedback and improvisational creation, traditional latency barriers in the "perception–adjustment" cycle are effectively removed, enabling dynamic optimization of the creative process. A standardized framework diagram delineates the core elements, logical relationships, and interaction pathways across all layers, providing a precise theoretical mapping for subsequent tool design.

2.3 Dialogue with existing theories

The triadic model advances targeted innovations that enable a substantive dialogue with major cognitive theories such as distributed cognition theory [17] and extended mind theory [18]. Through this dialogue, conceptual breakthroughs are achieved by addressing the unique characteristics of dance creation while remaining grounded in foundational principles of cognitive science. The model thus contributes a theoretical innovation situated at the intersection of technology, cognition, and artistic practice. The core differences and contributions of the triadic model, relative to existing theories, are summarized in Table 1.

Table 1. Core differences between the triadic model and existing theories

Comparison Dimension	Distributed Cognition Theory	Extended Mind Theory	Triadic Model
Primary focus	Distributed division of cognitive tasks	Extension of cognition through external tools	Co-evolution of technology, the body, and artistic expression
Role of the body	Execution substrate for cognitive tasks	Passive participant in cognitive processes	Principal agent in creative generation
Positioning of technology	Auxiliary node for cognitive tasks	External extension of cognition	Partner-like intermediary in co-creative processes
Artistic dimension	Not included in the core analytical framework	Emotional and cultural attributes de-emphasized	Emotional mapping and cultural symbolism are positioned as core components

In comparison with distributed cognition theory, both frameworks acknowledge the mediating role of technology within cognitive processes. However, distributed cognition theory emphasizes the distributed allocation of cognitive tasks across the “individual–tool–environment” system, treating the body merely as an execution substrate while overlooking the dominant role of bodily experience and affective motivation in dance creation. In contrast, the triadic model foregrounds the real-time coordination of “technology–body–environment,” situating the embodied cognition layer as the central hub and highlighting the body’s authority in data transformation and creative generation. This emphasis results in a theoretical construct that is more closely aligned with the embodied characteristics of dance composition. Relative to extended mind theory, both perspectives recognize the capacity of technology to augment cognitive abilities. Nevertheless, extended mind theory conceptualizes technology as an external cognitive tool and limits its concern to improvements in cognitive efficiency, thereby diminishing the emotional core and cultural attributes fundamental to artistic creation. The triadic model, through the explicit delineation of the artistic expression layer, incorporates emotional mapping and cultural symbolism as central components. The model asserts that technological augmentation must ultimately serve the conservation of artistic ontology, clarifying that the role of technology is to strengthen rather than replace the distinctive nature of artistic expression. This safeguards against the reduction of dance creation to a purely computational cognitive task. This theoretical dialogue and advancement provide a novel research perspective for the intersection of art and technology—one that simultaneously respects technological principles and preserves the essential characteristics of artistic practice.

3 INTELLIGENT TOOL DESIGN AND IMPLEMENTATION

3.1 Requirements analysis and design objectives

The requirements analysis for the tool was derived from a comprehensive investigation of dance composition scenarios and an extraction of core challenges, resulting in a three-dimensional requirements framework encompassing scenario adaptability, interaction experience, and functional extensibility. Scenario adaptability requirements emphasize the diversity of creative environments. The tool must support movement capture in standard indoor rehearsal studios, open outdoor spaces, and small temporary sites, while satisfying the need for unencumbered capture to avoid restricting choreographic bodily expression. Interaction experience requirements prioritize low-latency feedback, with interaction latency required to remain below 50 ms to accommodate real-time adjustments inherent to improvisational creation. Functional extensibility requirements include dual support for individual creative work and multi-user collaborative composition to address the coordination demands of different creative modes.

Based on these requirements, three core design objectives were established. First, the balance between technological precision and artistic freedom is pursued by constraining three-dimensional reconstruction error to within 2 cm and joint angle error to below 0.5°, while employing lightweight design to prevent operational burdens from disrupting the creative process. Second, personalized adaptation capability is implemented through the construction of dedicated feature libraries tailored to stylistic characteristics across dance domains—including the expressive bodily resonance of classical dance, the explosive dynamics of modern dance, and the highly stylized motion patterns of ethnic dance—while also providing tiered user interfaces for professional and non-professional dancers. Third, a scenario-based modular architecture is constructed, offering rapidly deployable modules for education, digital heritage preservation, and immersive performance. These modules respectively address the requirements of pedagogical feedback, high-fidelity recording, and real-time interactive engagement.

3.2 Technical architecture design

The tool adopts a four-layer closed-loop architecture comprising perception, processing, interaction, and application. These layers are integrated through standardized data interfaces to ensure efficient coordination and precise alignment between technical performance and creative requirements. The perception layer employs a multimodal sensing fusion scheme equipped with an MPU9250 nine-axis IMU and the MediaPipe v0.10.9 monocular vision system. Joint angle, center-of-mass trajectory, force application characteristics, and motion velocity data are synchronously captured at a sampling rate of 100 Hz. The IMU ensures local precision in limb movement, while the visual system provides global pose calibration. A spatiotemporal alignment algorithm is implemented to eliminate heterogeneity between the two data sources. The processing layer integrates edge computing with lightweight AI. Model deployment on the device side is implemented using TensorFlow Lite v2.15. An unscented Kalman filter is applied to mitigate data drift, with the process noise matrix Q set to $[1e^{-4}, 1e^{-4}, 1e^{-4}]$ and the observation noise matrix R set to $[1e^{-3}]$, ensuring stable dynamic capture. A lightweight Transformer model comprising four encoder layers and a 128-dimensional hidden layer is used to extract dance style features at 10 ms per frame, enabling real-time style adaptation.

The interaction layer is designed with a multimodal feedback mechanism. Force distribution patterns are visualized through heatmaps, haptic rhythm cues are transmitted via Bluetooth 5.3 Low Energy technology, and voice-based interaction delivers style optimization suggestions. Overall feedback latency is maintained below 38 ms. The application layer adopts a modular architecture. The composition management module supports movement library storage, version iteration, and export in BVH, FBX, and JSON formats. The collaborative creation module enables multi-terminal data synchronization using a 5G network and Redis caching, supporting concurrent creation for up to eight users. The scenario adaptation module enables rapid switching among educational, digital heritage, and performance modes, corresponding respectively to movement correction, high-fidelity archival recording, and real-time generation. Data transmission is implemented through Bluetooth 5.3 Low Energy for real-time streaming. Local data are stored using AES-256 encryption, and cloud backups are processed through anonymization protocols, balancing data security with accessibility.

Figure 1 presents the complete technical architecture of the mobile-interactive intelligent dance composition tool, encompassing the tool's client interface, the motion data relay service, and the dance interaction processing module. The modular composition and data flow logic of the four-layer architecture—perception, processing, interaction, and application—are clearly illustrated, providing an intuitive representation of the technical pathways underlying multimodal motion acquisition, lightweight style model loading, and related system functions.

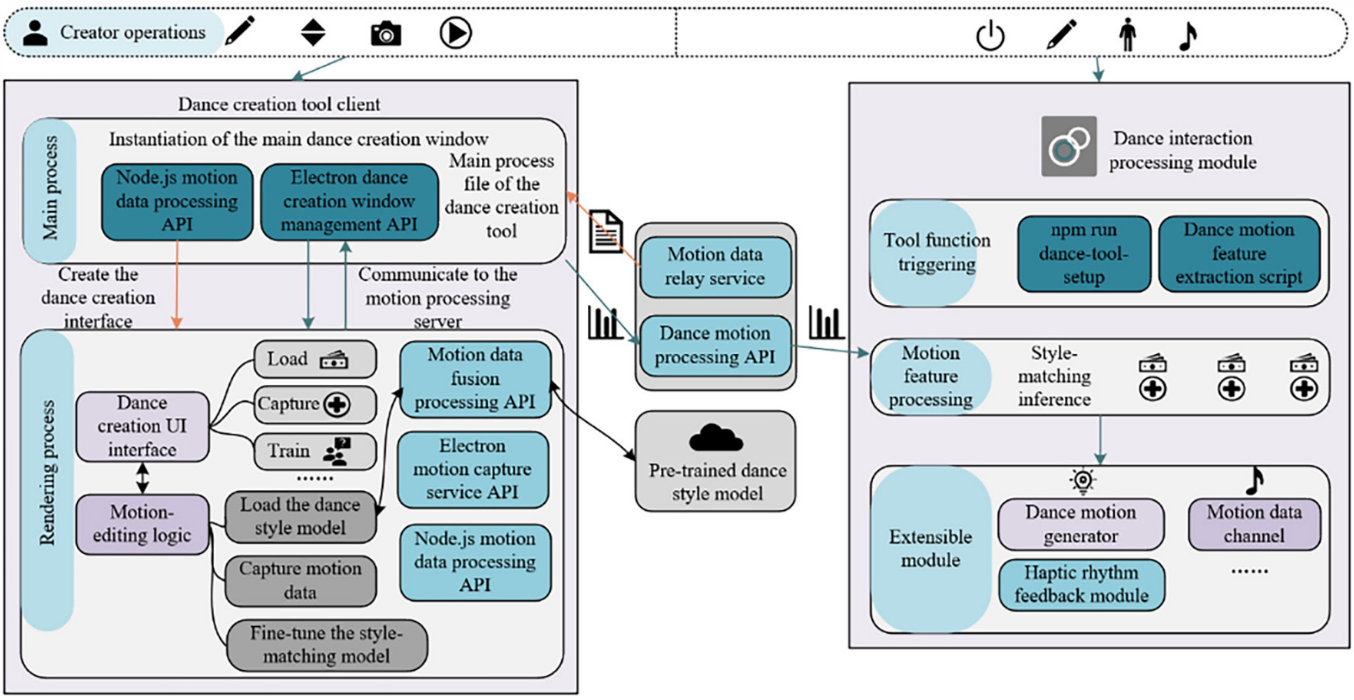


Fig. 1. Overview of the system architecture for the mobile-interactive intelligent dance composition tool

3.3 Implementation of core functional modules

Based on the four-layer technical architecture, four core functional modules were implemented. These modules operate through deeply integrated data pathways to support the full creative workflow of capture–generation–feedback–adaptation. Their key technologies, performance indicators, and application scenarios are summarized in Table 2. The mobile motion capture module functions as the primary data entry component and adopts a fusion scheme combining MediaPipe-based visual pose estimation with IMU data. The visual system outputs 2D joint keypoints, while the

IMU provides 3D motion pose data. A Perspective-*n*-Point (PnP) algorithm is applied to solve three-dimensional pose estimations. To address abnormal data resulting from motion blur or occlusion, a Random Sample Consensus (RANSAC) algorithm is introduced to remove outliers. The final reconstruction accuracy is maintained within 2 cm, satisfying precision requirements across multiple dance genres. The intelligent-assisted generation module is constructed using a lightweight Generative Adversarial Network (GAN). The encoder extracts semantic features from the input movements, and the decoder, combined with pre-trained style libraries for five major dance genres—classical, modern, ethnic, jazz, and contemporary—generates motion variants. A semantic consistency loss function is applied to ensure that generated variants remain aligned with the core semantics of the original movement, preventing unintended deviations in creative intent during style transfer.

The real-time feedback optimization module integrates biomechanical modeling with a dance style database. The biomechanical model computes joint loads and force efficiency based on the human skeletal–muscular structure, while the style database stores characteristic motion parameters representative of various dance genres. Their combined output yields three quantitative indicators—accuracy, coordination, and style congruence—along with corresponding visualized optimization suggestions. The scenario-adaptive module enables configuration-based switching across functional combinations. In the educational mode, emphasis is placed on real-time motion correction and instructional feedback, including numerical indicators of joint angle deviation. In the digital heritage mode, high sampling rate capture and multi-format archival functionality are activated to ensure long-term preservation value of motion data. In the performance mode, multi-user collaboration interfaces and real-time motion generation are optimized to support improvisational interactive creation in stage environments.

Table 2. Technical parameters and application scenarios of core functional modules

Module Name	Core Technologies	Performance Indicators	Primary Application Scenarios
Mobile motion capture module	MediaPipe–IMU fusion and RANSAC	3D reconstruction error ≤ 2 cm	Movement data acquisition across diverse environments
Intelligent-assisted generation module	Lightweight GAN and semantic consistency loss	Supports five dance style categories and generation latency ≤ 50 ms	Style transfer and motion variant generation
Real-time feedback optimization module	Biomechanical modeling and a dance style database	Three quantitative indicators and feedback latency ≤ 38 ms	Precision enhancement and style optimization of movements
Scenario-adaptive module	Modular functional configuration	Mode switching time ≤ 1 s	Education, heritage preservation, and immersive performance

3.4 Prototype validation and reproducibility

A three-tier validation framework was designed to assess the tool's technical effectiveness and methodological transparency, covering technical performance, data consistency, and user usability, while transparency was further supported through detailed reporting of the implementation and experimental procedures.

For technical performance evaluation, three representative creative environments were selected: an indoor rehearsal studio, an outdoor plaza, and a confined 10 m^2 space. Standardized movement sequences performed by 10 professional dancers were used as test samples, and the OptiTrack Prime 13 high-precision optical motion capture system served as the ground-truth reference. The results indicate that the average three-dimensional reconstruction errors across the three environments were 1.5 cm, 1.8 cm, and 1.7 cm, respectively; interaction latency remained stable between 32 and 38 ms; and the system sustained continuous operation for up to eight hours on a single charge, meeting the core performance requirements for mobile creative work. Reproducibility testing employed the publicly available Human3.6M and AIST++ dance motion datasets for cross-validation. The average capture error for the 15 action categories in Human3.6M was 1.6 cm, while the corresponding error for the five dance genres in AIST++ was 1.8 cm. Deviations from the benchmark optical system were consistently maintained within 0.3 cm, demonstrating performance stability across standardized datasets.

An initial usability assessment was conducted with 10 professional dancers and 10 non-professional dancers over a two-week pilot trial. Task completion time, operational error rate, and satisfaction scales were used as evaluation metrics. Professional users reported a satisfaction score of 8.2/10 for motion capture accuracy and style generation, while non-professional users rated the satisfaction of the operation workflow at 7.8/10. Based on this feedback, refinements were made to the interface's icon layout and feedback prompt mechanisms. The study provides detailed descriptions of the core methodological components, including the multimodal sensing configuration, data fusion workflow, model architecture, and experimental procedures. Technical aspects such as hardware configuration, software environment, parameter settings, and data preprocessing steps are explicitly reported to support methodological transparency. These descriptions establish a practical foundation for future implementation, comparison, and extension of the proposed approach by other researchers.

4 CREATIVE BEHAVIOR ANALYSIS: RESEARCH DESIGN

4.1 Research hypotheses

Based on the core logic of the triadic model of mobile-interactive dance composition and the technical characteristics of the tool, four progressive research hypotheses were formulated, establishing a systematic validation framework spanning tool effectiveness, behavioral evolution, group adaptability, and cognitive mechanisms. The four hypotheses are interrelated: Hypothesis 1 anchors the comparative advantages of the tool; Hypothesis 2 focuses on the dynamic behavior evolution; Hypothesis 3 examines the adaptability across groups; and Hypothesis 4 investigates the underlying cognitive mechanisms. Together, they form a complete validation chain encompassing tool performance–behavioral outcomes–cognitive characteristics. The formulation of the hypotheses is grounded in theoretical foundations and empirical evidence. Hypothesis 1 is derived from the tool's mobile adaptability and lightweight design, positing superior efficiency, flexibility, and lower cognitive load relative to Kinect and professional systems. Hypothesis 2 draws from embodied cognition theory regarding the reshaping effects of technology on cognition, emphasizing the co-evolution of data-driven and experiential elements. Hypothesis 3 is motivated by the group differences identified in the requirements analysis, highlighting the universal value of modular design. Hypothesis 4 incorporates cognitive psychology mechanisms linking feedback and immersion, introducing gaze-path entropy as a quantitative indicator of cognitive complexity.

4.2 Experimental design

To systematically validate the hypotheses, a mixed experimental design was implemented, combining controlled variables with repeated measures to ensure the reliability and validity of the results. Sixty creators participated in the study, divided into professional and non-professional groups with thirty participants each. The professional group included creators from classical, modern, and ethnic dance backgrounds, with ten participants from each style and more than five years of choreographic experience. The non-professional group consisted of individuals with over one year of dance-learning experience but no choreographic background. Age and gender ratios were matched across groups to control for extraneous variables. A $2 \times 3 \times 2$ mixed design was adopted. Tool type served as a between-subjects factor with three levels: the intelligent tool developed in this study, Kinect v2, and the OptiTrack professional system.

Each tool group contained 20 participants with matched proportions of professional and non-professional creators. Creative duration served as a within-subjects factor with two levels (short-term and long-term). Creator type was a between-subjects factor, forming an interaction structure with tool type. The experimental procedure consisted of short-term and long-term stages. The short-term experiment required participants to complete a three-minute dance-segment composition task, focusing on assessing immediate tool effectiveness. The long-term experiment spanned four weeks, during which participants engaged in weekly sessions to develop a sequence of four progressively advanced works, enabling the tracking of behavioral evolution. A unified thematic cue—natural imagery—was used to ensure comparability across groups. The short-term task required the creation of a single thematic segment, whereas the long-term task involved a series of creations that gradually deepened across stages of imagery extraction, movement design, and stylistic integration. This experimental design allowed between-subjects comparisons to capture differences across tool types, while within-subjects measurements revealed longitudinal evolution. When combined with analyses of group-level differences, the design fully addressed the validation needs of the four hypotheses. Task standardization and theme consistency further controlled for confounding variables such as task difficulty and creative motivation.

5 EXPERIMENTAL RESULTS AND ANALYSIS

Figure 2 illustrates the accuracy differences among the three motion capture tools across an indoor rehearsal studio, an outdoor plaza, and a confined 10 m^2 space. The intelligent tool developed in this study achieved accuracy errors between 1.5 and 1.8 cm across all environments, with a fluctuation range of only 0.3 cm, indicating stable performance under mobile conditions. Kinect v2 exhibited errors between 3.2 and 3.5 cm, maintaining consistent variability but demonstrating substantially lower precision overall. The OptiTrack professional system delivered the highest accuracy, with errors between 0.7 and 0.9 cm; however, its physical footprint and deployment requirements prevented operation in outdoor settings and small confined spaces. In terms of functional responsiveness, the proposed intelligent tool maintained an average core function latency of 32–38 ms, while synchronization latency during eight-user collaborative creation remained below 50 ms, meeting the requirements of real-time interactive composition. Kinect v2 exhibited latency in the range of 72–80 ms, and OptiTrack maintained latency between 18 and 22 ms. Performance validation results further demonstrated the technical stability of the proposed intelligent tool. Across the public datasets Human3.6M and AIST++, the average capture errors were 1.5 cm and 1.8 cm, respectively, with deviations from reference data remaining within 0.3 cm.

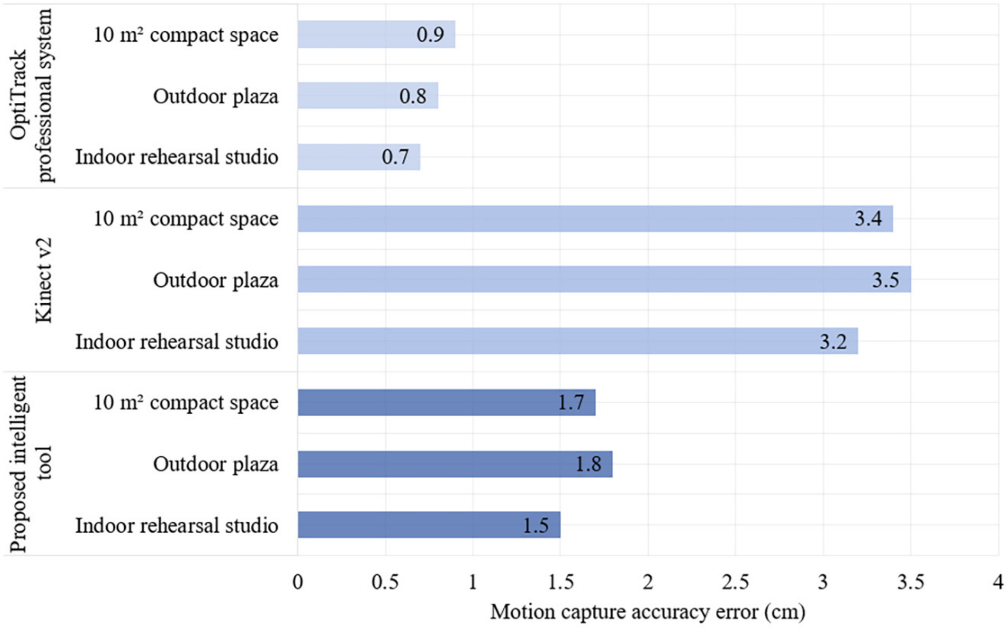


Fig. 2. Comparative accuracy error of motion capture tools across creative environments

Figure 3 demonstrates the dynamic evolution of creative behavior. For professional creators using the intelligent tool, the movement generation rate increased from 12 movements per minute in the short-term phase to 18 movements per minute by Week 4 of the long-term phase, while cognitive load (NASA–TLX score) decreased from 15 to 8. Non-professional creators exhibited a similar pattern, with movement generation rates increasing from 8 to 14 movements per minute and cognitive load decreasing from 18 to 12. In comparison, the Kinect v2 group showed only moderate improvements: professional creators' movement generation rates increased from 10 to 14 movements per minute, and cognitive load decreased from 17 to 14. These findings indicate that the intelligent tool more effectively facilitates a shift in creative cognition from experience dependence to a data–experience co-evolution model. Differences across creator types and dance genres were also observed. Professional creators demonstrated an 85% utilization rate of the style-matching module, compared with 60% among non-professional creators. Among dance genres, creators working in ethnic dance exhibited a 90% utilization rate of the cultural symbol extraction function, higher than that of classical dance (80%) and modern dance (75%). These patterns indicate that the tool's modular design effectively accommodates diverse functional needs across creator groups.

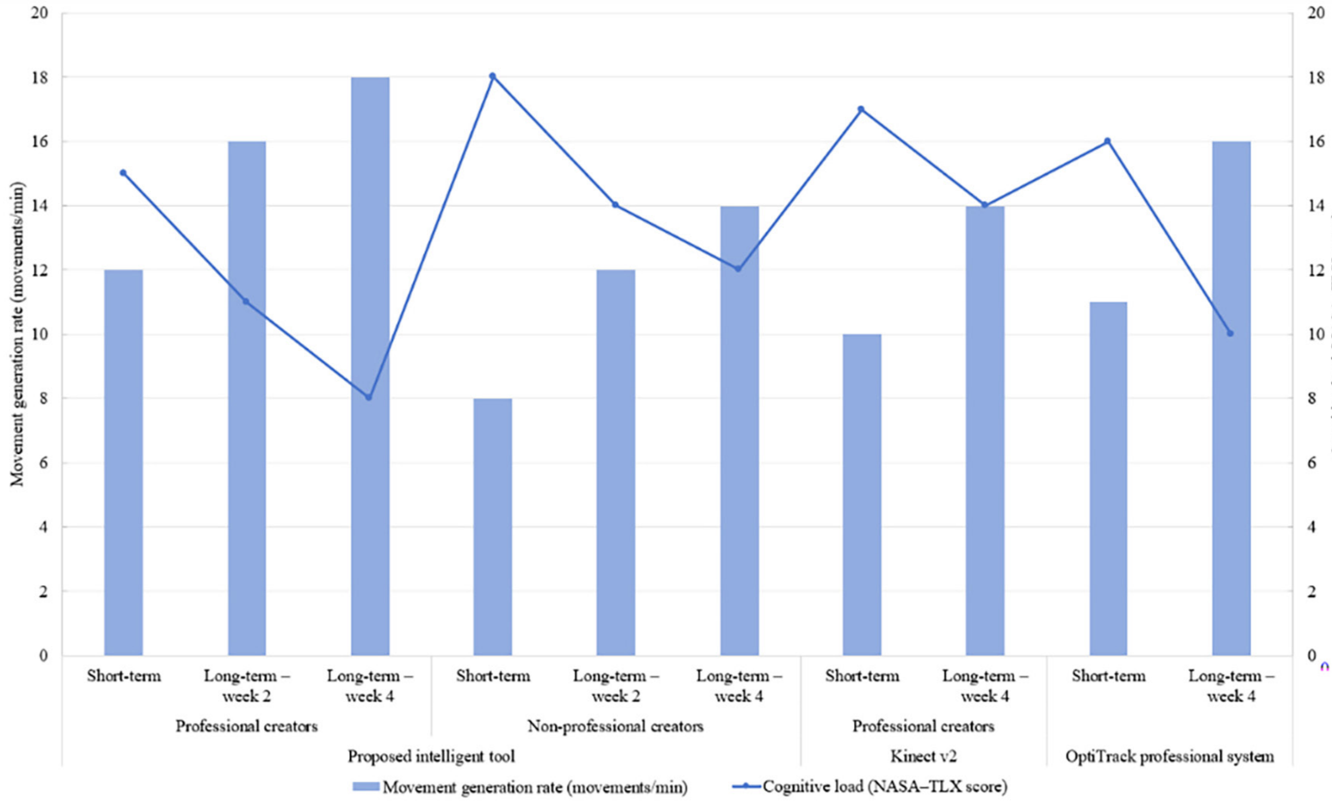


Fig. 3. Evolution of movement generation rate and cognitive load for professional and non-professional creators across tools

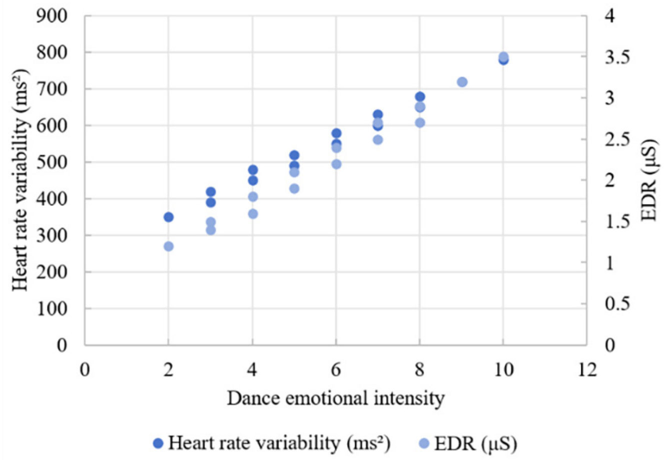


Fig. 4. Correlation distribution between dance emotional intensity and physiological indicators

As shown in Figure 4, a significant positive correlation was observed between emotion intensity and heart rate variability ($r = 0.68$, $p < 0.01$): as emotion intensity increased from 2 to 10, heart rate variability rose from 350 ms^2 to 780 ms^2. Electrodermal response (EDR) was also positively correlated with movement force ($r = 0.75$, $p < 0.01$), increasing from 1.2 μS to 3.5 μS as force ratings increased from 3 to 10. These results demonstrate the capacity of physiological indicators to serve as quantitative proxies for artistic expressive processes. Analysis of creative flow states showed that the intelligent tool group achieved a significantly higher flow score (8.2) than the Kinect v2 group (6.8), with no significant difference from the OptiTrack group (8.0). Creative flow scores were negatively correlated with cognitive load ($r = -0.72$, $p < 0.01$). Qualitative feedback further revealed that 90% of professional creators perceived no reduction in the uniqueness of artistic expression when using the tool, and 85% of non-professional creators reported that motion-assisted generation lowered the entry barrier to creative composition. These findings validate the tool's ability to balance technological augmentation with the preservation of artistic ontology.

Table 3. User experience and usability evaluation results

Evaluation Dimension/ Tool Type	Proposed Intelligent Tool (Professional Creators)	Proposed Intelligent Tool (Non-Professional Creators)	Kinect v2 (Professional Creators)	Kinect v2 (Non-Professional Creators)	OptiTrack Professional System (Professional Creators)
Tool learning time (min)	8	12	15	20	25
Operational satisfaction (1–5)	4.8	4.5	3.2	2.8	4.0
Task error rate (%)	5	8	12	18	6
Core function usage rate (%)	92	80	75	60	85
Operational smoothness (1–5)	4.7	4.3	3.0	2.5	4.2

To evaluate usability and creator acceptance under practical operational conditions, a user experience and usability assessment was conducted. As shown in Table 3, professional creators required only eight minutes to learn the intelligent tool, substantially shorter than the 15 minutes required for Kinect v2 and the 25 minutes required for the OptiTrack system. Non-professional users demonstrated similar advantages, requiring 12 minutes compared with 20 minutes for Kinect v2. Operational satisfaction scores for the intelligent tool reached 4.8 for professional creators and 4.5 for non-professional creators—significantly higher than the scores for Kinect v2 (3.2 and 2.8) and comparable to the OptiTrack system (4.0). Core function usage rate and operational smoothness exhibited similar superiority. These findings indicate that the tool's low-threshold design markedly enhanced acceptance across diverse creator groups. The usability advantages demonstrated relative to mainstream comparison tools provide experiential support for the tool's broader practical deployment.

6 CONCLUSION

This study addressed the central inquiry of integrating mobile interaction technologies with dance composition by constructing a triadic theoretical framework linking technology, the body, and cognition and by designing and implementing a mobile intelligent dance composition tool that integrates multimodal sensing with lightweight AI. Multidimensional experimental validation demonstrated the tool's technical effectiveness, creative value, and adaptability across diverse scenarios. As shown by the results, stable motion capture accuracy and low latency responsiveness were maintained across multiple creative environments; creative efficiency was substantially improved for both professional and non-professional creators;

and creative cognition was shifted from experience dependence toward a data–experience co-evolution model. Furthermore, the system achieved a balance between technological augmentation and the preservation of artistic expressiveness. Its practical effectiveness in dance education and digital heritage preservation further confirmed its potential to address core challenges within the field. This study fills a theoretical gap at the intersection of embodied cognition, mobile intelligence, and artistic creation, while also providing a transparent methodological framework for future implementation and comparison that may serve as a reference for subsequent interdisciplinary research on mobile-intelligent creative practices.

7 ACKNOWLEDGEMENTS

This paper was funded by the Key Project of the Ministry of Education in the "14th Five-Year Plan" for National Educational Research: Innovative Research on Ethnic Dance Teaching in Colleges and Universities (Grant No.: JYKY234208); and the 2025 Heze Social Science Planning Project "Research on the Ecological Aesthetic Imagery of the Shang Yang Dance" (Grant No.: 2025-ZZ-35).

8 REFERENCES

[1] X. Wang, Y. Han, C. Wang, Q. Zhao, X. Chen, and M. Chen, "In-Edge AI: Intelligentizing mobile edge computing, caching and communication by federated learning," *IEEE Network*, vol. 33, no. 5, pp. 156–165, 2019. https://doi.org/10.1109/MNET.2019.1800286

[2] L. Liu, "The impact of mobile applications on personalized learning paths in dance education," *International Journal of Interactive Mobile Technologies*, vol. 19, no. 5, pp. 128–143, 2025. https://doi.org/10.3991/ijim.v19i05.54525

[3] M. del Carmen Rodríguez-Hernández and S. Ilarri, "AI-based mobile context-aware recommender systems from an information management perspective: Progress and directions," *Knowledge-Based Systems*, vol. 215, p. 106740, 2021. https://doi.org/10.1016/j.knosys.2021.106740

[4] K. Chen, Y. Zu, and D. Wang, "Design and implementation of intelligent creation platform based on artificial intelligence technology," *Journal of Computational Methods in Science and Engineering*, vol. 20, no. 4, pp. 1109–1126, 2020. https://doi.org/10.3233/JCM-204240

[5] S. K. Jha, B. H. Kumari, K. Anand, J. Kanjalkar, R. B. Gaddam, and S. Kumar, "Intelligent task prediction and partial computation offloading in mobile edge cloud computing," *International Journal of Interactive Mobile Technologies*, vol. 19, no. 18, pp. 106–117, 2025. https://doi.org/10.3991/ijim.v19i18.57233

[6] S. M. C. Palpán and L. G. Castro, "Autonomy and dialogue during creation: A qualitative study on the experiences of creation in dance of a group of students," *Artseduca*, vol. 27, pp. 72–87, 2020.

[7] M. G. Valladares Gonzalez, "Artistic dance creation: One experience of interdisciplinary research," *Tercio Creciente*, no. 14, pp. 37–48, 2018.

[8] A. E. Yankovskaya and V. B. Obukhovskaya, "Foundations of creation of a complex of applied intelligent systems for diagnostics of psychological safety and cognitive sphere of patients with a neurological pathology," *Pattern Recognition and Image Analysis*, vol. 30, no. 4, pp. 741–747, 2020. https://doi.org/10.1134/S1054661820040252

[9] M. R. Nogueira, J. B. Simões, J. M. de Carvalho, and P. Menezes, "'Move in Tempo': Involving the audience through their movement in installation art," *International Journal of Arts and Technology*, vol. 15, no. 5, pp. 1–22, 2025. https://doi.org/10.1504/IJART.2025.146787

[10] L. Vuarnesson *et al*., "Shared diminished reality: A new VR framework for the study of embodied intersubjectivity," *Frontiers in Virtual Reality*, vol. 2, p. 646930, 2021. https://doi.org/10.3389/frvir.2021.646930
[11] K. M. Darda and E. S. Cross, "The computer, A choreographer? Aesthetic responses to randomly-generated dance choreography by a computer," *Heliyon*, vol. 9, no. 1, p. e12750, 2023. https://doi.org/10.1016/j.heliyon.2022.e12750
[12] B. Paulchamy, A. Yahya, N. Chinnasamy, and K. Kasilingam, "Facial expression recognition through transfer learning: Integration of VGG16, ResNet, and AlexNet with a multiclass classifier," *Acadlore Transactions on AI and Machine Learning*, vol. 4, no. 1, pp. 25–39, 2025. https://doi.org/10.56578/ataiml040103
[13] R. Namane, E. Boutellaa, S. E. Salem, and Y. Babaci, "A residual temporal convolutional with attention neural network for electromyogram-based hand gesture recognition," *International Journal of Computational Methods and Experimental Measurements*, vol. 13, no. 3, pp. 739–748, 2025. https://doi.org/10.56578/ijcmem130320
[14] A. O. Shabalina, "Two meanings of 'Cognition' in the theory of embodied cognition," *Tomsk State University Journal*, vol. 463, pp. 69–72, 2021. https://doi.org/10.17223/15617793/463/9
[15] M. Yang, J. Amankwah-Amoah, and H. Chunjia, "Exploring human-machine collaboration in metaverse communities: Product vs. service focus in enhancing user immersion," *Transportation Research Part E: Logistics and Transportation Review*, vol. 204, p. 104423, 2025. https://doi.org/10.1016/j.tre.2025.104423
[16] G. Baudoux, "The benefits and challenges of artificial intelligence image generators for architectural ideation: Study of an alternative human-machine co-creation exchange based on sketch recognition," *International Journal of Architectural Computing*, vol. 22, no. 2, pp. 201–215, 2024. https://doi.org/10.1177/14780771241253438
[17] Z. Liu, N. Nersessian, and J. Stasko, "Distributed cognition as a theoretical framework for information visualization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1173–1180, 2008. https://doi.org/10.1109/TVCG.2008.121
[18] M. Bernini, "Supersizing narrative theory: On intention, material agency, and extended mind-workers," *Style*, vol. 48, no. 3, pp. 349–366, 2014. https://doi.org/10.5325/style.48.3.0349

9 AUTHOR

Yanjun Jiang graduated from Jiangsu Normal University in 2016 and currently serves as a faculty member at the School of Music and Dance, Heze University, Heze 274015, China, specializing in research on Chinese ethnic folk dances and the appreciation of Chinese and foreign dance works (E-mail: wlxjjj@126.com).

Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 4 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i04.58589

PAPER

Beyond Tech-Fluent Generations: Investigating Cross-Generational Technology Adoption Patterns in Collaborative Online Learning Spaces

Shamim Akhter[1] (✉), **Rabindra Dev Prasad**[1], **Mengqiu Tan**[2], **Sehrish Iftikhar**[3]

[1]INTI International University, Nilai, Malaysia

[2]Guangdong University of Petrochemical Technology, Maoming, China

[3]University of Southern Punjab, Multan, Pakistan

shamim.akhter@newinti.edu.my

ABSTRACT

This study challenges the prevailing digital native's paradigm by examining technology adoption patterns across different generational cohorts in collaborative online learning environments. It investigates how generational differences influence technology acceptance, usage behaviors, and learning outcomes in digital educational spaces. A mixed-methods approach was employed, combining quantitative surveys (n = 847) and qualitative interviews (n = 32) across four generational cohorts: Generation Z (born 1997–2012), Millennials (1981–1996), Generation X (1965–1980), and Baby Boomers (1946–1964). The study utilized the extended technology acceptance model (TAM2) framework, incorporating social influence and cognitive instrumental processes. Findings reveal significant variations in technology adoption patterns that transcend traditional generational assumptions. While Generation Z demonstrated higher initial technology acceptance rates (M = 4.23, SD = 0.87), Generation X showed superior sustained engagement in collaborative learning activities (M = 4.45, SD = 0.76). Baby Boomers exhibited unexpected adaptability when provided with appropriate scaffolding and support mechanisms. The digital natives concept oversimplifies technology adoption behaviors. Cross-generational collaboration in online learning spaces benefits from differentiated instructional design approaches that acknowledge varying technological competencies while leveraging the unique strengths of each generational cohort.

KEYWORDS

digital natives, technology adoption, learning opportunities, online collaboration, educational technology

1 INTRODUCTION

The concept of digital natives, first introduced by Prensky [1], has dominated educational technology discourse for over two decades. This paradigm suggests

Akhter, S., Prasad, R. D., Tan, M., Iftikhar, S. (2026). Beyond Tech-Fluent Generations: Investigating Cross-Generational Technology Adoption Patterns in Collaborative Online Learning Spaces. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(4), pp. 75–89. https://doi.org/10.3991/ijim.v20i04.58589

Article submitted 2025-09-12. Revision uploaded 2025-12-12. Final acceptance 2025-12-12.

that individuals born into the digital age possess innate technological competencies that fundamentally differ from previous generations, labeled as digital immigrants. However, recent scholarship has begun to question this binary classification, arguing that it oversimplifies the complex relationship between age, technology adoption, and learning preferences [2], [3]. The rapid acceleration of online learning, particularly following the COVID-19 pandemic, has created unprecedented opportunities to observe cross-generational technology adoption patterns in educational contexts [4]. Contemporary online learning environments increasingly feature learners from multiple generational cohorts, creating dynamic spaces where different technological perspectives and competencies intersect [5].

This study addresses a critical gap in current literature by examining how different generational cohorts adapt to and utilize collaborative online learning technologies. Rather than accepting the digital native's framework as definitive, this research investigates the nuanced ways in which age-related factors influence technology adoption, usage patterns, and collaborative learning outcomes in digital educational environments. The significance of this research extends beyond theoretical considerations. As educational institutions increasingly adopt blended and fully online learning modalities, understanding cross-generational technology adoption patterns becomes essential for designing inclusive and effective digital learning experiences [6]. The findings of this study have implications for instructional design, technology implementation strategies, and the development of age-inclusive online learning environments.

2 LITERATURE REVIEW

2.1 The digital natives paradigm: Evolution and critique

The digital natives concept emerged from observations that individuals born after 1980 appeared to interact with technology differently than previous generations [1]. This framework suggested that exposure to digital technologies from an early age created neuro-plastic changes that fundamentally altered learning preferences and cognitive processing patterns. Digital natives were characterized as multitasking, visually oriented learners who preferred interactive and immediate learning experiences. However, empirical research has increasingly challenged these assumptions. Thompson [7] found no significant differences in multitasking abilities between digital natives and digital immigrants in academic contexts. Similarly, Margaryan et al. [8] demonstrated that university students, despite being classified as digital natives, exhibited limited technological competencies beyond basic social media and communication tools.

Recent meta-analyses have further undermined the digital native's paradigm. Gallardo-Echenique et al. [9] analyzed 58 studies and found inconsistent evidence supporting generational differences in technology use and learning preferences. The authors concluded that individual factors such as socioeconomic status, educational background, and personal motivation were more predictive of technology adoption than generational membership. Thapa et al. [10] show that emotional intelligence influences student engagement with mobile technologies.

2.2 Technology acceptance in educational contexts

The TAM, developed by Davis [11] provides a theoretical framework for understanding individual technology adoption decisions. The model identifies perceived usefulness and perceived ease of use as primary determinants of technology acceptance. Subsequent extensions, including TAM2 [12] and the Unified Theory of Acceptance and Use of Technology (UTAUT) [13], have incorporated additional factors such as social influence, facilitating conditions, and individual characteristics.

In educational contexts, technology acceptance models have been adapted to account for the unique characteristics of learning environments. Al-Emran et al. [14] found that self-efficacy, social influence, and institutional support significantly influenced technology acceptance among university students. However, most studies in this domain have focused on homogeneous age groups, limiting our understanding of cross-generational technology adoption patterns. Recent bibliometric analyses of mobile learning adoption research confirm the predominant focus on single-cohort studies [34].

2.3 Collaborative online learning across generations

Collaborative online learning environments present unique challenges and opportunities for cross-generational interaction. Social constructivist theories suggest that diverse perspectives enhance learning outcomes through the negotiation of meaning and knowledge construction [15]. In online contexts, this diversity can be manifested through different technological competencies, learning preferences, and communication styles. Research demonstrates that social presence significantly influences learning satisfaction and persistence in online learning environments [33]. Research on age-diverse online learning teams has produced mixed findings. Some studies suggest that generational diversity can lead to enhanced problem-solving and creativity [16], with evidence showing that social media integration can enhance critical thinking in online learning contexts [35]. Conversely, other research indicates that technological skill disparities can create barriers to effective collaboration [17].

The COVID-19 pandemic provided an unprecedented natural experiment in cross-generational online learning adoption. Studies conducted during this period revealed that age-related technology adoption patterns were more complex than previously assumed. Park et al. [18] found that older learners demonstrated remarkable adaptability when provided with appropriate support systems, while younger learners sometimes struggled with the sustained focus required for online learning.

2.4 Research gaps and study rationale

Despite growing interest in cross-generational technology adoption, several research gaps remain. First, most existing studies focus on single generational cohorts rather than examining patterns across multiple age groups simultaneously. Second, most of the research has been conducted in formal educational settings, limiting our understanding of technology adoption in collaborative online learning environments. Third, existing literature often treats generational membership as a binary variable (digital native vs. digital immigrant) rather than examining the continuous nature of age-related technology adoption. Finally, few studies have employed mixed-methods approaches that capture both quantitative patterns and qualitative experiences of cross-generational technology adoption.

This study addresses these gaps by examining technology adoption patterns across four distinct generational cohorts in collaborative online learning environments, employing both quantitative and qualitative methodologies to provide a comprehensive understanding of cross-generational technology adoption behaviors.

3 METHODOLOGY

3.1 Research design

This study employed a concurrent mixed-methods design (Creswell and Plano Clark, 2017) to investigate cross-generational technology adoption patterns in collaborative online learning spaces. The quantitative component utilized a cross-sectional survey design to examine technology acceptance patterns across generational cohorts, while the qualitative component employed semi-structured interviews to explore individual experiences and perspectives.

3.2 Participants

The study recruited 847 participants from various online learning platforms and educational institutions across North America and Europe between January and September 2024. Participants were categorized into four generational cohorts based on birth year: Generation Z (n = 234, ages 12–27), Millennials (n = 298, ages 28–43), Generation X (n = 201, ages 44–59), and Baby Boomers (n = 114, ages 60–78).

Inclusion criteria required participants to have engaged in at least one collaborative online learning experience within the previous six months. Collaborative learning was defined as structured educational activities involving two or more participants working together toward shared learning objectives through digital platforms.

For the qualitative component, 32 participants were purposively selected from the survey respondents to ensure representation across all generational cohorts and various online learning contexts. The interview sample included eight participants from each generational cohort, with equal gender representation where possible.

3.3 Instruments

Technology adoption survey. The quantitative instrument was adapted from the TAM2 framework (Venkatesh and Davis, 2000) and included measures of perceived usefulness (α = 0.89), perceived ease of use (α = 0.92), social influence (α = 0.87), and behavioral intention (α = 0.91). Additional scales measured collaborative learning effectiveness (α = 0.88) and technology self-efficacy (α = 0.94). The survey also included demographic items and questions about specific online learning platforms used, frequency of collaborative activities, and preferred communication modalities. All items utilized 5-point Likert scales ranging from 1 (strongly disagree) to 5 (strongly agree).

The Technology Adoption Survey was developed based on the extended TAM2 framework (Venkatesh and Davis, 2000) and adapted specifically for collaborative online learning contexts.

Semi-structured interview protocol. The interview protocol explored participants' experiences with online learning technologies, collaboration strategies, challenges encountered, and adaptive behaviors. The following key questions are included:

Personal technology adoption journey. Technology adoption is a highly individualized process that unfolds through exposure, exploration, integration, and mastery phases. Generation Z participants described seamless progression from gaming and social apps to educational technologies, relying on intuitive navigation and peer-to-peer learning through online communities rather than formal instruction. Millennials experienced more structured adoption, beginning with email and basic internet during their education years and systematically expanding their skills as new platforms emerged. Their journeys involved deliberate skill-building phases, often driven by professional or educational requirements. Generation X participants took pragmatic, purpose-driven approaches, typically starting their technological journey in workplace environments where necessity drove rapid skill acquisition. They preferred intensive learning periods with familiar tools, adopting new technologies only when clear benefits were evident, and emphasized understanding underlying principles over interface memorization.

Baby Boomers initially showed resistance but gradually accepted and often enthusiastically embraced technologies that demonstrated clear value. Their adoption was frequently motivated by family connections, professional needs, or personal interests, with many describing feelings of accomplishment and empowerment as they mastered platforms that enabled social connections and lifelong learning.

Experiences with cross-generational collaboration. Cross-generational collaboration challenged stereotypes about age-related tech competencies, with younger participants surprised by older collaborators' adaptability and superior organizational skills, systematic problem-solving, and attention to detail that enhanced group productivity. Millennials often served as technological bridges, translating between communication styles and facilitating knowledge transfer in both directions, while Generation Z developed patience and thorough work habits by observing older peers' careful planning and quality control approaches. Generation X emphasized valuable mentoring relationships where they shared professional expertise and project management skills while learning emerging technologies and digital communication norms from younger collaborators.

Baby Boomers found collaboration initially intimidating but ultimately rewarding, discovering new platforms through younger participants' patient technical support while contributing valued analytical thinking, writing skills, and subject matter expertise. Cross-generational teams consistently produced higher-quality outcomes than age-homogeneous groups, combining technological fluency, life experience, and diverse perspectives to create comprehensive solutions and demonstrate that age-related learning differences were matters of approach rather than fundamental capability.

Perceived advantages and disadvantages of different technologies. Participants across generational cohorts identified distinct advantages and disadvantages of collaborative learning technologies, with priorities varying based on technological backgrounds and learning preferences. Video conferencing was universally appreciated for face-to-face interaction, particularly by older participants who valued the familiar meeting format, though Generation Z experienced more fatigue and preferred asynchronous communication for flexibility. Learning Management Systems received mixed reviews, with Generation X and Baby Boomers appreciating structured organization while younger participants criticized them as outdated and preferred streamlined, mobile-optimized interfaces similar to social media. Collaborative document editing and project management tools were generally well-received, though younger participants favored real-time editing and immediate feedback while older participants preferred structured revision processes with clear version control. Social learning platforms revealed generational divides, with older participants preferring threaded discussions for in-depth exchanges while Generation Z favored dynamic, multimedia-rich platforms supporting shorter, more frequent visual interactions.

Strategies for overcoming technological barriers. Participants developed diverse strategies for overcoming technological barriers, with approaches varying significantly based on generational cohorts, prior experience, and individual learning preferences. Generation Z typically employed trial-and-error exploration supplemented by peer consultation through social networks, demonstrating high tolerance for confusion while discovering features through serendipitous navigation, though they struggled with sustained troubleshooting requiring methodical diagnosis. Millennials combined systematic research with peer consultation, creating personal reference materials and developing hybrid approaches that mixed formal learning with informal experimentation, often serving as technological intermediaries between different experience levels. Generation X preferred systematic, methodical approaches focused on understanding underlying principles rather than memorizing steps, consulting multiple authoritative sources, and establishing relationships with technologically experienced colleagues for guidance. Baby Boomers developed patient, incremental strategies beginning with basic functionality, creating detailed personal reference materials, practicing privately before collaborative use, and benefiting from structured peer learning with other older adults who shared similar challenges.

Preferences for learning and communication modalities. Learning and communication modality preferences revealed complex patterns reflecting both generational influences and individual characteristics, challenging simplistic assumptions about age-related technology preferences while highlighting the need for multiple options to accommodate diverse learner needs. Generation Z demonstrated strong preferences for multimedia-rich, interactive experiences with video content and immediate feedback, favoring rapid, informal exchanges through multiple simultaneous channels but sometimes struggling with sustained engagement requiring extended focused attention. Millennials exhibited preferences for balanced approaches combining structured content with interactive elements, appreciating professional-grade platforms supporting both formal and informal styles with flexible cross-device access to accommodate complex scheduling demands. Generation X preferred clear structure and logical progression with opportunities for deep engagement, emphasizing quality over quantity in communication and appreciating connections between new information and existing professional knowledge. Baby Boomers demonstrated preferences for learning modalities providing clear guidance and comprehensive support, emphasizing personal connection and relationship-building through face-to-face interaction while valuing thoughtful, well-considered communication that acknowledged their life experience and expertise.

3.4 Data collection procedures

Quantitative data were collected through online surveys distributed via educational platform partnerships and social media channels. The survey required approximately 15–20 minutes to complete and was available in English, Spanish, and French.

Qualitative interviews were conducted via video conferencing platforms (Zoom or Microsoft Teams) based on participant preference. Interviews lasted 45–60 minutes and were audio-recorded with participant consent. All interviews were conducted by trained researchers using standardized protocols.

3.5 Data analysis

Quantitative data were analyzed using SPSS 28.0. Descriptive statistics characterized the sample and examined variable distributions. One-way ANOVA tested for differences across generational cohorts, with post-hoc Tukey tests identifying specific group differences. Structural equation modeling (SEM) using AMOS 26.0 examined relationships between variables within the extended TAM framework.

Qualitative data were analyzed using thematic analysis following Braun and Clarke's (2006) six-phase approach. Interviews were transcribed verbatim and coded independently by two researchers. Initial codes were organized into themes through iterative analysis and discussion. NVivo 12 software facilitated data organization and coding consistency checks.

3.6 Ethical considerations

The study received institutional review board approval from the lead researcher's institution (IRB #2023–114). All participants provided informed consent, and data were collected and stored following established privacy and confidentiality protocols. Participants were informed of their right to withdraw at any time without penalty.

4 RESULTS

4.1 Participant demographics

The final sample included 847 participants with representation across all target generational cohorts. Generation Z participants ($n = 234$, 27.6%) were predominantly students in higher education, while Millennials ($n = 298$, 35.2%) represented a mix of graduate students and early-career professionals. Generation X participants ($n = 201$, 23.7%) were primarily working professionals engaged in continuing education, and Baby Boomers ($n = 114$, 13.5%) included both retirees and late-career professionals.

The sample was 58.2% female, 40.1% male, and 1.7% non-binary or preferred not to specify. Educational attainment varied across cohorts, with higher percentages of advanced degrees among older participants. Technology access was generally high across all groups, with 96.3% reporting reliable internet access and personal computing devices.

4.2 Cross-generational technology adoption patterns

Technology acceptance variables. Significant differences emerged across generational cohorts for several technology acceptance variables. Generation Z demonstrated the highest mean scores for perceived ease of use ($M = 4.23$, $SD = 0.87$), followed by Millennials ($M = 4.01$, $SD = 0.93$), Generation X ($M = 3.78$, $SD = 1.02$), and Baby Boomers ($M = 3.45$, $SD = 1.15$). ANOVA results confirmed significant differences [$F(3,843) = 28.74$, $p < 0.001$]. Interestingly, perceived usefulness scores showed less dramatic generational variation. While Generation Z scored highest ($M = 4.31$, $SD = 0.79$), the differences between Millennials ($M = 4.18$, $SD = 0.84$), Generation X ($M = 4.22$, $SD = 0.81$), and Baby Boomers ($M = 4.09$, $SD = 0.89$) were not statistically significant [$F(3,843) = 2.89$, $p = 0.035$].

Social influence played a more prominent role for older participants. Baby Boomers reported the highest social influence scores ($M = 3.87$, $SD = 0.98$), significantly higher than Generation Z ($M = 3.42$, $SD = 1.04$) [$t(346) = 3.67$, $p < 0.001$].

Collaborative learning effectiveness. Contrary to digital natives' assumptions, Generation X participants reported the highest collaborative learning effectiveness scores ($M = 4.45$, $SD = 0.76$), followed by Baby Boomers ($M = 4.32$, $SD = 0.82$), Millennials ($M = 4.18$, $SD = 0.89$), and Generation Z ($M = 4.05$, $SD = 0.95$). These differences were statistically significant [$F(3,843) = 12.43$, $p < 0.001$].

Post-hoc analyses revealed that Generation X scored significantly higher than both Millennials and Generation Z, while Baby Boomers scored significantly higher than Generation Z. No significant difference was found between Generation X and baby boomers.

4.3 Platform usage patterns

Analysis of platform preferences revealed distinct generational patterns. Generation Z participants showed strong preferences for video-based platforms (85.2%) and social learning applications (73.4%). Millennials demonstrated more balanced usage across platform types, with high adoption of learning management systems (78.5%) and collaboration tools (71.8%).

Generation X participants showed the highest sustained engagement with traditional learning management systems (89.6%) but also embraced video conferencing tools (82.1%) for synchronous collaboration. Baby Boomers displayed more selective technology adoption, with strong preferences for email-based communication (91.2%) and structured discussion forums (76.3%).

4.4 Qualitative themes

Adaptive learning strategies. Across all generational cohorts, participants described developing adaptive strategies to navigate technological challenges. Generation Z participants often served as informal technology mentors but sometimes struggled with sustained attention in asynchronous learning environments. One Generation Z participant noted, "I can pick up new apps really quickly, but sometimes I miss the depth that comes from really focusing on one thing for a long time. The older people in my study group are actually better at that."

Generation X and baby boomer participants described systematic approaches to technology adoption, often investing more time in initial learning but demonstrating higher long-term retention and utilization. A Baby Boomer participant explained, "I may take longer to learn a new platform, but once I understand it, I tend to use it more thoroughly than some of the younger participants who seem to skim the surface."

Cross-generational collaboration benefits. Participants across all cohorts identified benefits from cross-generational collaboration. Younger participants valued the focus and analytical skills of older learners, while older participants appreciated the technological flexibility and creative approaches of younger cohorts.

A millennial participant observed, "The best project teams I've been on had people from different age groups. The younger people brought energy and technical skills, the middle-aged people brought project management, and the older people brought wisdom and perspective."

Technology barriers and solutions. Different generational cohorts faced distinct technological barriers. Generation Z participants reported challenges with sustained engagement and preferences for rapid feedback mechanisms. Millennials described difficulty balancing multiple technological platforms and information sources. Generation X participants identified time constraints as a primary barrier, preferring efficient, purpose-driven technology implementations. Baby boomers emphasized the importance of technical support and gradual implementation strategies.

4.5 Structural equation modeling results

The extended TAM model demonstrated good fit across all generational cohorts [χ^2/df = 2.34, CFI = 0.96, RMSEA = 0.041]. However, path coefficients varied significantly across groups.

For Generation Z, perceived ease of use was the strongest predictor of behavioral intention ($\beta = 0.67$, $p < 0.001$), while social influence showed minimal impact ($\beta = 0.12$, $p = 0.184$). Conversely, for Baby Boomers, both perceived usefulness ($\beta = 0.58$, $p < 0.001$) and social influence ($\beta = 0.41$, $p < 0.001$) significantly predicted behavioral intention.

Generation X and millennial patterns fell between these extremes, with moderate influences from all TAM variables. Notably, technology self-efficacy emerged as a significant mediator for all groups but was particularly important for Generation X ($\beta = 0.49$, $p < 0.001$) and Baby Boomers ($\beta = 0.52$, $p < 0.001$).

4.6 Longitudinal engagement patterns

Analysis of sustained engagement over the six-month study period revealed interesting generational differences. While Generation Z participants showed high initial adoption rates, their engagement declined significantly over time (initial M = 4.23, final M = 3.78, $p < 0.001$).

In contrast, Baby Boomer participants demonstrated increased engagement over time (initial M = 3.45, final M = 3.89, $p < 0.01$), suggesting that their initial hesitancy was overcome through experience and support. Generation X and millennial participants maintained relatively stable engagement levels throughout the study period, with Generation X showing slight increases in collaborative learning effectiveness scores over time.

5 DISCUSSION

5.1 Challenging the digital natives paradigm

The findings of this study provide compelling evidence that the digital natives paradigm oversimplifies cross-generational technology adoption patterns. While Generation Z participants demonstrated higher initial technology acceptance rates, their sustained engagement and collaborative learning effectiveness scores were lower than those of older cohorts. This pattern suggests that technological fluency, as commonly conceptualized, may not directly translate to effective learning outcomes in collaborative online environments. The ability to quickly navigate new technologies, while valuable, appears to be distinct from the skills required for sustained, productive online collaboration.

Generation X participants' superior performance in collaborative learning effectiveness metrics challenges assumptions about age-related technological competence. Their systematic approach to technology adoption, combined with strong project management and analytical skills, appeared to compensate for any initial technological learning curves.

5.2 The role of social influence and support systems

The significantly higher impact of social influence on older participants' technology adoption decisions highlights the importance of peer support and institutional backing for successful technology implementation. This finding is consistent with recent research demonstrating that older adults' technology adoption is significantly influenced by social cognitive factors, particularly observational learning and peer support mechanisms [19]. Studies have shown that peer-to-peer community learning environments, including the use of "super-users" as technology champions, effectively support sustained technology use among older adults [20]. Furthermore, research based on social cognitive theory has demonstrated that social influence mechanisms, including vicarious learning through peer observation, significantly enhance self-efficacy and technology adoption rates among older adults [21].

5.3 Implications for instructional design

The research findings have significant implications for instructional design in online learning environments. Rather than assuming homogeneous technological competencies based on age, designers should consider differentiated approaches that acknowledge varying strengths across generational cohorts. This approach aligns with established principles of differentiated instruction, which emphasize adapting content, process, and assessment to accommodate diverse learner needs while maintaining consistent learning objectives [22]. Research in adult learning contexts demonstrates that differentiated instructional approaches, which recognize individual differences in readiness, learning preferences, and technological backgrounds, significantly enhance engagement and learning outcomes across diverse populations [23]. For Generation Z learners, instructional designs might incorporate more frequent feedback mechanisms, varied interaction modalities, and explicit scaffolding for sustained engagement. Millennial learners might benefit from streamlined technology ecosystems that reduce platform switching and cognitive load. Generation X learners appear to thrive with clear structure and purpose-driven technology implementations, while Baby Boomer learners benefit from gradual introduction strategies and robust technical support systems.

These differentiated strategies reflect research showing that successful technology training programs for diverse age groups incorporate multiple learning modalities, personalized learning paths, and flexible instructional approaches that respect individual learning preferences [24].

5.4 The value of cross-generational learning communities

Perhaps the most significant finding of this study is the evidence for the value of cross-generational learning communities. Rather than segregating learners by age or assumed technological competence, the research suggests that diverse age groups can enhance learning outcomes through complementary skills and perspectives. This finding corroborates emerging research on cross-generational collaboration, which demonstrates that age-diverse teams exhibit improved creativity, problem-solving capabilities, and decision-making effectiveness compared to age-homogeneous groups [25]. Studies of cross-generational workplace learning have documented that multigenerational collaboration fosters innovation through the integration of diverse perspectives, with younger generations contributing technological fluency while older generations provide strategic thinking and contextual expertise [26].

The qualitative data revealed numerous examples of successful cross-generational collaboration, where different technological approaches and learning strategies combined to create more robust learning outcomes than would be achieved by age-homogeneous groups. Research on intergenerational learning in digital environments has shown that structured opportunities for cross-generational interaction, including peer mentoring and collaborative projects, break down age-related stereotypes while leveraging the unique strengths of each generation [27]. Furthermore, studies demonstrate that creating psychologically safe environments where learners from different generations feel valued and heard is essential for effective cross-generational knowledge sharing [28].

5.5 Technology adoption as a continuous process

The longitudinal analysis revealed that technology adoption is not a static characteristic but rather an ongoing process that varies significantly across individuals and contexts. The finding that Baby Boomer participants increased their engagement over time while Generation Z participants decreased suggests that initial adoption patterns may not predict long-term success. This pattern aligns with longitudinal research on technology adoption, which demonstrates that sustained usage behavior is often distinct from initial adoption, with early usage patterns fortifying long-term engagement trajectories [29]. Recent studies of AI chatbot adoption in higher education have revealed similar patterns, showing significant declines in usage behavior among initially enthusiastic adopters over extended periods, emphasizing the importance of examining temporal dynamics rather than relying on single time-point assessments [30]. Research on sustained technology engagement indicates that factors promoting initial adoption (such as perceived ease of use) may differ from factors supporting long-term usage (such as perceived usefulness and social support), particularly across different age groups [31].

This pattern has implications for both research methodology and practical implementation. Single time-point assessments of technology adoption may not capture the dynamic nature of learning and adaptation in online environments. Studies examining digital engagement of older adults have emphasized that barriers and facilitators differ across engagement stages (nonuse, initial adoption, and sustained use), highlighting the need for longitudinal approaches to understand the complete spectrum of technology adoption experiences [32].

5.6 Limitations and future research directions

Several limitations should be acknowledged. First, the study focused on participants with existing access to technology and internet connectivity, potentially excluding populations with limited technological resources. Second, the generational cohort definitions, while widely accepted, may not capture individual variation within age groups. Third, the study was conducted primarily in North American and European contexts, limiting generalizability to other cultural and educational contexts. Finally, the six-month follow-up period may not be sufficient to capture long-term technology adoption patterns.

Future research should explore technology adoption patterns in diverse cultural contexts, examine the role of socioeconomic factors in cross-generational learning, and investigate longer-term impacts of age-diverse online learning communities.

5.7 Practical implications for educational institutions

The findings suggest several practical recommendations for educational institutions implementing online learning programs. First, technology training and support should be tailored to different generational cohorts' needs and preferences rather than employing one-size-fits-all approaches. Second, cross-generational learning opportunities should be actively promoted and supported through structured collaboration activities and mentoring programs. Third, technology selection and implementation should consider the diverse needs and competencies of multi-generational learner populations.

Finally, ongoing support and adaptation strategies should be implemented to address the dynamic nature of technology adoption and learning engagement over time.

6 CONCLUSIONS

This study provides compelling evidence that the digital natives paradigm is insufficient for understanding technology adoption patterns in contemporary online learning environments, revealing that cross-generational technology adoption is far more nuanced than previously assumed. The finding that Generation X participants achieved the highest collaborative learning effectiveness scores, despite lower initial technology acceptance ratings, challenges fundamental assumptions about age and technological competence, while evidence that Baby Boomer participants can successfully adapt to new technologies with appropriate support undermines stereotypes about older learners' capabilities. Most importantly, this research demonstrates the value of cross-generational learning communities, suggesting that educational institutions and online learning designers should recognize generational differences as assets that enhance learning outcomes through complementary perspectives and skills. Moving forward, the field of educational technology must move beyond simplistic generational categories toward more nuanced understanding of individual differences, learning preferences, and adaptive capabilities, developing age-inclusive approaches that leverage the strengths of all generational cohorts to create effective and equitable digital learning environments.

7 GENERATIVE AI DECLARATION

No generative AI tools were used in the drafting or revision of this manuscript.

8 FUNDING

Not applicable.

9 CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

10 DATA AVAILABILITY STATEMENT

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request, subject to ethical approval and participant consent requirements.

11 REFERENCES

[1] M. Prensky, "Digital natives, digital immigrants Part 1," *On the Horizon*, vol. 9, no. 5, pp. 1–6, 2001. https://doi.org/10.1108/10748120110424816

[2] P. A. Kirschner and P. De Bruyckere, "The myths of the digital native and the multitasker," *Teaching and Teacher Education*, vol. 67, pp. 135–142, 2017. https://doi.org/10.1016/j.tate.2017.06.001

[3] L. Reid, D. Button, and M. Brommeyer, "Challenging the myth of the digital native: A narrative review," *Nursing Reports*, vol. 13, no. 2, pp. 573–600, 2023.

[4] C. Hodges, S. Moore, B. Lockee, T. Trust, and A. Bond, "The difference between emergency remote teaching and online learning," *EDUCAUSE Review*, vol. 55, no. 3, pp. 12–23, 2020.

[5] Z. Wang, L. Chen, and T. Anderson, "Post-pandemic online learning: Generational differences and adaptive strategies," *Distance Education*, vol. 44, no. 2, pp. 234–252, 2023. https://doi.org/10.1080/01587919.2023.2198765

[6] L. Chen and M. Zhang, "Inclusive design principles for multi-generational online learning environments," *Educational Technology Research and Development*, vol. 72, no. 3, pp. 445–468, 2024. https://doi.org/10.1007/s11423-024-10298-7

[7] P. Thompson, "The digital natives as learners: Technology use patterns and approaches to learning," *Computers & Education*, vol. 65, pp. 12–33, 2013. https://doi.org/10.1016/j.compedu.2012.12.022

[8] A. Margaryan, A. Littlejohn, and G. Vojt, "Are digital natives a myth or reality? University students' use of digital technologies," *Computers & Education*, vol. 56, no. 2, pp. 429–440, 2011. https://doi.org/10.1016/j.compedu.2010.09.004

[9] E. E. Gallardo-Echenique, L. Marqués-Molías, M. Bullen, and J. W. Strijbos, "Let's talk about digital learners in the digital era," *The International Review of Research in Open and Distributed Learning*, vol. 16, no. 3, pp. 156–187, 2015. https://doi.org/10.19173/irrodl.v16i3.2196

[10] P. P. Thapa, N. M. Zayed, M. N. Alam, V. S. Nitsenko, S. Rudenko, and D. Svyrydenko, "Mediating and moderating role of emotional intelligence between mobile phone use and affective commitment among undergraduate students in academic institutes," *Current Psychology*, vol. 44, pp. 6610–6626, 2025. https://doi.org/10.1007/s12144-025-07661-x

[11] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, vol. 13, no. 3, pp. 319–340, 1989. https://doi.org/10.2307/249008

[12] V. Venkatesh and F. D. Davis, "A theoretical extension of the technology acceptance model: Four longitudinal field studies," *Management Science*, vol. 46, no. 2, pp. 186–204, 2000. https://doi.org/10.1287/mnsc.46.2.186.11926

[13] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," *MIS Quarterly*, vol. 27, no. 3, pp. 425–478, 2003. https://doi.org/10.2307/30036540

[14] M. Al-Emran, V. Mezhuyev, and A. Kamaludin, "Technology acceptance model in M-learning context: A systematic review," *Computers & Education*, vol. 125, pp. 389–412, 2018. https://doi.org/10.1016/j.compedu.2018.06.008

[15] L. S. Vygotsky, *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press, 1978.

[16] J. M. Trujillo-Torres, I. Aznar-Díaz, M. P. Cáceres-Reche, T. Mentado-Labao, and A. Barrera-Corominas, "Intergenerational learning and its impact on the improvement of educational processes," *Educ. Sci.,* vol. 13, no. 10, p. 1019, 2023. https://doi.org/10.3390/educsci13101019

[17] Q. Tang, S. Kamarudin, S. N. A. Rahman, and X. Zhang, "Bridging gaps in online learning: A systematic literature review on the digital divide," *J. Educ. Learn.,* vol. 14, no. 1, pp. 161–176, 2025.

[18] R. Sultana and R. Palaroan, "COVID-19 and transition to online learning: Evidence from a Sino-Foreign university in China," *J. Instr. Pedagog.,* vol. 28, pp. 1–20, 2023. https://www.aabri.com/manuscripts/213501.pdf

[19] T. K. Hof and S. L. van Hooijdonk, "Creating a social learning environment for and by older adults in the use and adoption of smartphone technology to age in place," *Frontiers in Public Health,* vol. 9, p. 568822, 2021. https://doi.org/10.3389/fpubh.2021.568822

[20] M. K. Peeters, J. van Grondelle, and K. L. Winters, "Promoting technology adoption and engagement in aging," in *Connected Health in Aging Adults*, New York, NY: Springer Publishing, 2023, pp. 23–45.

[21] H. Chen, X. Zhang, and Y. Wang, "Digital technologies-enhanced older adults health management: Developing a five-dimensional extension of social learning theory for community settings," *Frontiers in Public Health*, vol. 13, p. 1627983, 2025. https://doi.org/10.3389/fpubh.2025.1627983

[22] C. A. Tomlinson, "Differentiated instruction in response to academically diverse classrooms," in *International Guide to Student Achievement*, J. Hattie and E. M. Anderman, Eds., New York, NY: Routledge, 2013, pp. 275–277.

[23] C. Y. Colbert, C. M. Foshee, A. Prelosky-Leeson, M. Schleicher, and R. King, "Differentiated instruction as a viable framework for meeting the needs of diverse adult learners in health professions education," *Med. Sci. Educ.,* vol. 33, no. 4, pp. 975–984, 2023. https://doi.org/10.1007/s40670-023-01808-w

[24] M. L. Anderson and K. R. Thompson, "How to train and engage a multigenerational workforce," TalentLMS Research Report, 2025. [Online]. Available: https://www.talentlms.com/research/multigenerational-workforce

[25] S. A. Lima and M. M. Rahman, "Generational diversity and inclusion: HRM challenges and opportunities in multigenerational workforces," *SSRN Electron. J.*, 2025. https://ssrn.com/abstract=5330092

[26] K. Rodriguez and S. Kim, "Advanced learning strategies for multigenerational workforce," *eLearning Industry Journal*, vol. 18, no. 1, pp. 45–62, 2024.

[27] J. M. Trujillo-Torres, I. Aznar-Díaz, M. P. Cáceres-Reche, T. Mentado-Labao, and A. Barrera-Corominas, "Intergenerational learning and its impact on the improvement of educational processes," *Educ. Sci.*, vol. 13, no. 10, art. 1019, 2023. https://doi.org/10.3390/educsci13101019

[28] S. E. Nursyamsi, N. Siregar, and S. R. Suminar, "The effectiveness of digital communication in bridging the generational gap in the workplace," *J. Law Soc. Sci. Humanit.*, vol. 2, no. 1, pp. 172–187, 2024. https://myjournal.or.id/index.php/JLSSH/article/view/270

[29] V. Venkatesh and M. G. Morris, "A longitudinal field investigation of gender differences in individual technology adoption decision-making processes," *Organizational Behavior and Human Decision Processes*, vol. 83, no. 1, pp. 33–60, 2000. https://doi.org/10.1006/obhd.2000.2896

[30] M. Skjuve, I. A. Sætre, A. Følstad, and P. B. Brandtzaeg, "A longitudinal study on artificial intelligence adoption: Understanding the drivers of ChatGPT usage behavior change in higher education," *Frontiers in Artificial Intelligence*, vol. 6, p. 1324398, 2023. https://doi.org/10.3389/frai.2023.1324398

[31] S. Taylor and P. Todd, "Understanding information technology usage: A test of competing models," *Information Systems Research*, vol. 6, no. 2, pp. 144–176, 1995. https://doi.org/10.1287/isre.6.2.144

[32] J. Y. Lee, H. R. Yoon, and K. J. Kim, "Digital engagement of older adults: Scoping review," *Journal of Medical Internet Research*, vol. 24, no. 12, p. e40192, 2022. https://doi.org/10.2196/40192

[33] Y. Zhang, H. B. Mohamed, and D. Liu, "A systematic literature review on the relationships between social presence, learning satisfaction and persistence in online learning," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 18, no. 5, pp. 44–61, 2024. https://doi.org/10.3991/ijim.v18i05.47799

[34] Y. Singh and P. K. Suri, "A bibliometric analysis of the literature on mobile learning adoption and continuance in the field of education," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 17, no. 17, pp. 38–58, 2023. https://doi.org/10.3991/ijim.v17i17.40965

[35] N. D. Abd Halim, M. A. Mutalib, N. M. Zaid, M. Mokhtar, F. A. Majid, and W. N. Elia Haslee Sharil, "The use of social media to enhance critical thinking in online learning among higher education students," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 18, no. 6, pp. 56–66, 2024. https://doi.org/10.3991/ijim.v18i06.48033

12 AUTHORS

Shamim Akhter is with the INTI International University, Persiaran Perdana BBN, Putra Nilai, 71800 Nilai, Negeri Sembilan, Malaysia (E-mail: shamim.akhter@newinti.edu.my).

Rabindra Dev Prasad is with the INTI International University, Persiaran Perdana BBN, Putra Nilai, 71800 Nilai, Negeri Sembilan, Malaysia (E-mail: rabindra.prasad@newinti.edu.my).

Mengqiu Tan is with the School of Marxism, Guangdong University of Petrochemical Technology, Maoming 525000, China (E-mail: m.tan@gdupt.edu.cn).

Sehrish Iftikhar is a Lecturer at the University of Southern Punjab, Multan, Pakistan (E-mail: sehrishiftikhar@usp.edu.pk).

Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 4 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i04.60517

PAPER

An Intelligent Feedback Mechanism for Mobile English Learning Based on Learning Behavior Analytics

Jingyi Cai(✉)

Xuchang Vocational Technical College, Xuchang, China

caijingyi1984@163.com

ABSTRACT

The rapid proliferation of mobile learning technologies has reshaped English learning contexts, giving rise to fragmented learning patterns and increasing demands for personalization. However, feedback mechanisms in current mobile English learning systems remain overly focused on knowledge error correction, with limited consideration of learners' cognitive processes, thereby constraining learning effectiveness. The deep integration of learning behavior analytics with intelligent feedback is regarded as a promising pathway to address this limitation. This study aims to address a central research question: how can fine-grained learning behavior analytics be leveraged to integrate observable behavioral data with latent metacognitive states in order to construct a pedagogically adaptive and personalized intelligent feedback mechanism for mobile English learning? To this end, a multi-task deep knowledge tracing (DKT) approach incorporating metacognitive assessment was proposed. On this basis, a closed-loop framework integrating data perception, joint modeling, and multidimensional feedback was constructed, with optimization strategies tailored to mobile learning scenarios. Experimental results based on real-world mobile learning data demonstrated that the proposed approach significantly outperformed conventional DKT models in both knowledge tracing accuracy and metacognitive state recognition accuracy. Moreover, the resulting intelligent feedback mechanism effectively enhanced learners' English learning performance and metacognitive abilities. This study extends existing language learning theory through the integration of metacognitive modeling with knowledge tracing, introduces a discipline-adaptive multi-task intelligent feedback modeling paradigm, and provides a practical pathway for the intelligent enhancement of mobile English learning systems.

KEYWORDS

mobile English learning, learning behavior analytics, metacognitive assessment, deep knowledge tracing (DKT), intelligent feedback mechanism, multi-task learning

1 INTRODUCTION

The widespread adoption of mobile terminals and the continuous evolution of mobile Internet technologies have profoundly redefined the temporal and

Cai, J. (2026). An Intelligent Feedback Mechanism for Mobile English Learning Based on Learning Behavior Analytics. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(4), pp. 90–104. https://doi.org/10.3991/ijim.v20i04.60517

Article submitted 2025-11-07. Revision uploaded 2026-01-08. Final acceptance 2026-01-08.

spatial boundaries of English learning. As a result, fragmentation and ubiquity have emerged as defining characteristics of contemporary mobile English learning environments [1–3]. In parallel, the deep integration of intelligent technologies—such as artificial intelligence and big data—with the educational domain [4, 5] has reached a global consensus as a critical driver of digital education transformation, thereby offering technological opportunities to address the long-standing limitations of personalization in traditional language education. Despite these advances, substantial deficiencies remain in the feedback mechanisms employed by current mobile English learning platforms. First, feedback content has tended to be highly homogenized, rendering it insufficiently responsive to individual learners' cognitive differences [6, 7]. Second, feedback logic has been predominantly restricted to correctness judgments at the knowledge level, while the underlying cognitive processes and psychological states reflected in learning behaviors have largely been neglected. Third, feedback delivery has shown limited adaptability to fragmented mobile learning contexts, with inadequate lightweight design and insufficient immediacy, thereby constraining its practical effectiveness [8, 9].

Against this backdrop, the introduction of learning behavior analytics and metacognitive theory has been widely recognized as providing essential support for optimizing intelligent feedback mechanisms [10–12]. Learning behavior data encompass not only observable indicators—such as response accuracy and interaction operations—but also latent dimensions, including confidence perception and strategy selection, enabling a more precise characterization of learners' authentic learning states [13, 14]. Moreover, metacognitive interventions have been demonstrated to effectively guide learners' self-monitoring and self-regulation processes, playing a critical role in enhancing learner autonomy and learning effectiveness in language acquisition contexts [15, 16]. The systematic integration of learning behavior analytics and metacognitive modeling therefore establishes a solid foundation for the development of intelligent feedback mechanisms with enhanced pedagogical value in mobile English learning environments.

Building on the foregoing context, four core research questions are addressed in this study: (a) how a multidimensional learning behavior and metacognitive data acquisition system can be constructed to accommodate mobile English learning scenarios; (b) how metacognitive assessment can be integrated with DKT to enable accurate joint modeling of learners' knowledge states and cognitive characteristics in English learning; (c) how a three-dimensional intelligent feedback mechanism—encompassing knowledge error correction, metacognitive guidance, and learning behavior strategy optimization—can be designed based on precise modeling outcomes; and (d) how the effectiveness of the proposed intelligent feedback mechanism in improving English learning performance and metacognitive ability can be empirically validated. In response to these questions, the primary objective of this study is defined as the construction of a behavior-cognition-metacognition integrated intelligent feedback mechanism for mobile English learning. Through this framework, feedback content is intended to be fully personalized, feedback logic is pedagogically grounded, and feedback presentation is contextually adapted, thereby providing learners with comprehensive and multidimensional learning support.

The remainder of this study is organized below. Section 2 elaborates on the core theoretical foundations underpinning the research and establishes the overall research framework. Section 3 designs and implements a multidimensional learning behavior and metacognitive data acquisition system tailored to mobile learning contexts. Section 4 proposes a multi-task DKT model incorporating metacognitive assessment, with detailed descriptions of the model architecture and training optimization procedures. Section 5 constructs a three-dimensional intelligent feedback mechanism based on model outputs and optimizes feedback presentation and delivery strategies for mobile scenarios. Section 6 validates the performance of the proposed model and the effectiveness of the intelligent feedback mechanism through controlled experiments. The final section summarizes the principal findings and conclusions.

2 RESEARCH FRAMEWORK

A closed-loop research framework consisting of data acquisition-joint modeling-feedback generation-effect validation was established in this study. The constituent modules are tightly interconnected and iteratively refined through bidirectional interactions, thereby ensuring the precision, adaptability, and effectiveness of the intelligent feedback mechanism. The data acquisition module is designed to collect multidimensional data aligned with mobile English learning scenarios. Inputs include learners' diverse learning behaviors on mobile platforms as well as metacognition-related information. Through data cleaning and feature engineering procedures, standardized datasets are produced as outputs, providing high-quality data support for subsequent modeling processes. The joint modeling module is driven by a multi-task DKT model incorporating metacognitive assessment. Standardized data generated by the data acquisition module are taken as inputs. Through a multi-task learning architecture, knowledge state tracing and metacognitive trait identification are simultaneously performed. Outputs consist of precise diagnostic results, including learners' knowledge weaknesses, metacognitive state categories, and behavioral strategy tendencies.

The feedback generation module constructs three-dimensional intelligent feedback content based on the diagnostic outcomes of the joint modeling module while optimizing feedback presentation and delivery strategies for mobile learning contexts. Inputs comprise the diagnostic results produced by joint modeling, whereas outputs include personalized feedback targeting knowledge error correction, metacognitive guidance, and learning behavior strategy optimization. Through this process, feedback content is ensured to align closely with learners' needs, and feedback forms are adapted to fragmented mobile learning scenarios. The effect validation module evaluates the effectiveness of the intelligent feedback mechanism through controlled experiments. Data on learning performance, metacognitive ability, and user experience are collected after the deployment of the feedback mechanism. These evaluation data serve as inputs, and empirical conclusions regarding feedback effectiveness are generated as outputs. Issues identified during validation are subsequently fed back into the data acquisition and joint modeling modules, enabling iterative optimization of the overall framework. Within the entire framework, the MetaKD-DKT model serves as the core driving component. Its capability for joint diagnosis of knowledge and metacognition functions as a critical bridge between data acquisition and feedback generation, ensuring that the intelligent feedback mechanism achieves its central objectives of personalization, pedagogical grounding, and contextual adaptation through fine-grained learning behavior analytics.

3 CONSTRUCTION OF THE MOBILE ENGLISH LEARNING BEHAVIOR AND METACOGNITIVE DATA SYSTEM

Data acquisition is designed with the primary objective of supporting the joint assessment of knowledge states and metacognition. The construction of the data system adheres to three core principles—multidimensionality, quantifiability, and contextual adaptability—to ensure that learners' learning states and cognitive traits can be comprehensively and accurately characterized. The data system encompasses three categories of core data. Basic profile data, including learners' age, English proficiency level, and learning goals, are collected to provide a foundational basis for subsequent personalization. Observable learning behavior data comprise response sequences, mobile interaction behaviors, fragmented learning characteristics, and device-assisted behaviors. Specifically, response sequences include item identifiers, correctness, and response time; mobile interaction behaviors involve touch duration, swipe frequency, and mis-touch rollback actions; fragmented learning characteristics cover learning time slots, session duration, and interruption frequency; and device-assisted behaviors include voice reading and dictionary usage. Metacognitive data consist of response confidence, problem-solving traces, and post-task reflective behaviors. Response confidence is categorized into four levels—guessing, uncertain, relatively certain, and highly certain. Problem-solving traces include hint viewing and option switching, whereas post-task reflective behaviors encompass incorrect-item marking, note taking, and willingness to reattempt practice.

Data acquisition is implemented through a combination of automated logging and supplementary survey-based collection. The primary acquisition tool is an event-logging plug-in developed using the Flutter framework and embedded within the mobile English learning application. This plug-in enables real-time and automated collection of observable learning behavior data, thereby ensuring objectivity and temporal accuracy. For metacognitive information that is difficult to capture automatically, supplementary online questionnaires are employed to enhance the completeness of metacognitive data. Data storage adopts a hybrid strategy integrating local caching and cloud synchronization. SQLite is used to support temporary local data caching, while cloud synchronization is implemented through Alibaba Cloud Object Storage Service (OSS). This design effectively mitigates data loss risks caused by unstable mobile network conditions and ensures the integrity and security of the collected data.

4 DESIGN OF THE METAKD-DKT MODEL INCORPORATING METACOGNITIVE ASSESSMENT

The MetaKD-DKT model is designed around the dual core objectives of knowledge tracing and metacognitive assessment, thereby overcoming the limitation of conventional knowledge tracing models that focus exclusively on knowledge states. This design is grounded in the intrinsic association between learning behaviors and metacognition. By simultaneously capturing learners' levels of knowledge mastery and metacognitive characteristics, a comprehensive and precise diagnosis of learner states is achieved. A multitask learning architecture is adopted, in which collaborative training of the primary and auxiliary tasks enables the model to acquire discriminative metacognitive features while learning the core capability of knowledge tracing. As a result, model generalization performance and diagnostic accuracy are effectively enhanced, providing high-precision data support for the subsequent construction of a three-dimensional intelligent feedback mechanism. The overall architecture of the MetaKD-DKT model incorporating metacognitive assessment is illustrated in Figure 1.

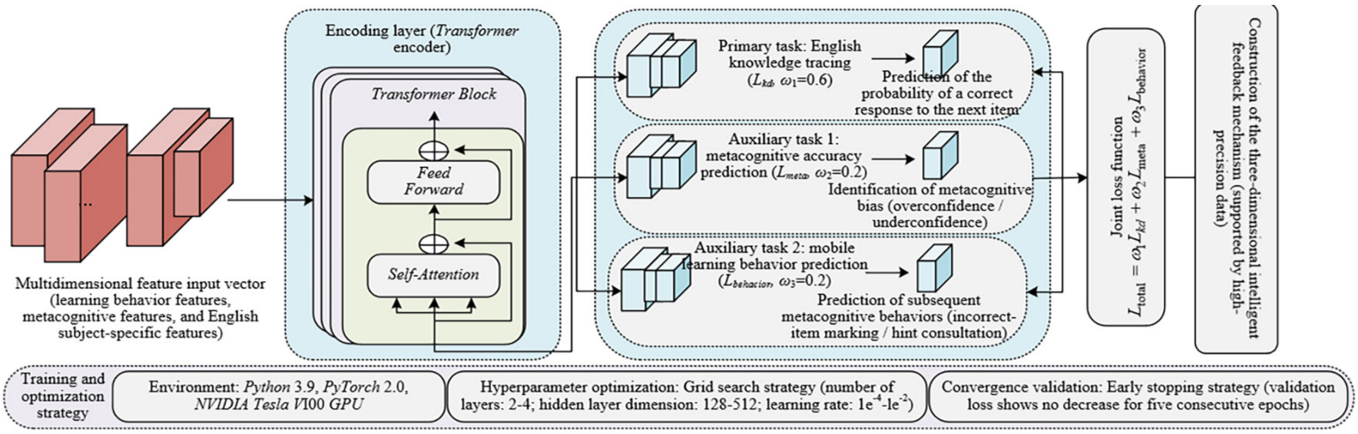


Fig. 1. Architecture of the MetaKD-DKT model incorporating metacognitive assessment

The model architecture consists of an encoding layer and a multi-task learning layer, enabling multidimensional feature fusion and multi-objective joint optimization. The encoding layer employs a Transformer encoder, whose self-attention mechanism is more effective than traditional structures based on long short-term memory (LSTM) in capturing long-range dependencies in sequential learning behaviors. Inputs are provided in the form of multidimensional vectors integrating learning behavior features, metacognitive features, and English subject-specific features. After encoding, historical state representations embedding comprehensive information are generated. The multi-task learning layer comprises one primary task and two auxiliary tasks. The primary task is English knowledge tracing, with the objective of predicting the probability of a learner's correct response to the next item, thereby enabling precise identification of knowledge weaknesses. Auxiliary Task 1 focuses on metacognitive accuracy prediction, in which metacognitive biases—such as overconfidence or underconfidence—are identified by evaluating the alignment between learners' confidence levels and their actual knowledge mastery. Auxiliary Task 2 targets mobile learning behavior prediction, with emphasis placed on anticipating learners' subsequent metacognitive behaviors, including incorrect-item marking and hint consultation. To balance training priorities across tasks, a weighted loss function is adopted, which is defined as:

$$L_{total} = \omega_1 L_{kd} + \omega_2 L_{meta} + \omega_3 L_{behavior} \tag{1}$$

where, L_{kd} denotes the loss associated with the knowledge tracing task, with $\omega_1 = 0.6$. The term L_{meta} represents the loss for metacognitive accuracy prediction, whereas $L_{behavior}$ corresponds to the loss for mobile learning behavior prediction, with $\omega_2 = \omega_3 = 0.2$.

Model training was implemented using Python 3.9 and the PyTorch 2.0 framework. Parallel acceleration was achieved with an NVIDIA Tesla V100 GPU, thereby ensuring computational efficiency during training. Hyperparameter optimization was conducted via a grid search strategy, with primary attention devoted to key parameters of the Transformer encoder. The search space included 2–4 encoder layers, hidden layer dimensions ranging from 128 to 512, and learning rates spanning from $1e^{-4}$ to $1e^{-2}$. Optimal hyperparameter configurations were selected through exhaustive traversal of parameter combinations in conjunction with validation set performance. Model convergence was assessed by continuous monitoring of the validation loss curve. When no decrease in validation loss was observed for five consecutive epochs, an early stopping strategy was triggered to terminate training. This procedure can effectively mitigate the risk of overfitting and ensure robust generalization performance on unseen data.

5 CONSTRUCTION OF A THREE-DIMENSIONAL INTELLIGENT FEEDBACK MECHANISM FOR MOBILE ENGLISH LEARNING BASED ON METAKD-DKT

The design of the three-dimensional intelligent feedback mechanism is guided by three core principles: pedagogical adaptability, personalization, and contextual adaptability. These principles are intended to ensure the educational value, precision, and practical usability of the feedback. Under the principle of pedagogical adaptability, feedback content is required to closely align with the instructional logic of English language education and to focus on the core objectives of language knowledge construction and learning ability development. Consequently, feedback is designed not only to perform error correction but also to provide instructional guidance. The principle of personalization is grounded in the diagnostic outputs of the MetaKD-DKT model. Differentiated feedback content is generated based on learners' identified knowledge weaknesses and metacognitive characteristics, thereby avoiding homogenized feedback delivery. The principle of contextual adaptability emphasizes the fragmented nature of mobile learning scenarios. Accordingly, lightweight and multimodal feedback forms are adopted to balance informational completeness with contextual convenience, ultimately enhancing feedback reception efficiency and learners' willingness to apply the feedback.

The generation of three-dimensional feedback content is driven by the high-precision diagnostic results produced by the MetaKD-DKT model. Through a quantitative matching mechanism, diagnostic features are accurately mapped to feedback content, forming three mutually coordinated feedback dimensions: knowledge-level feedback, metacognitive feedback, and learning behavior strategy feedback. To quantify the degree of correspondence between model diagnostic results and feedback templates, a feedback matching score model is introduced and defined as follows:

$$S_f = \sum_{i=1}^{n} \omega_i \cdot \text{sim}(F_i, T_i) \tag{2}$$

where, S_f denotes the feedback matching score. The term F_i represents the i-th category of diagnostic features output by the MetaKD-DKT model, such as identified knowledge weaknesses, types of metacognitive bias, and behavioral tendencies. The term T_i corresponds to the feature representation of the feedback template in the same dimension. The function sim( ) denotes the cosine similarity, which is used to compute the degree of correspondence between features. The weight ω_i reflects the relative importance of the i-th diagnostic feature, satisfying $\sum_{i=1}^{n} \omega_i = 1$. Weight allocation is determined using the analytic hierarchy process (AHP) based on the contribution of each feature to learning improvement.

At the knowledge level, feedback is centered on the knowledge weaknesses predicted by the model. Optimal knowledge correction templates are selected according to the feedback matching score, and targeted feedback content is delivered to directly remediate language knowledge gaps. At the metacognitive level, which constitutes the core innovation of the proposed mechanism, feedback is generated based on the model's metacognitive accuracy assessment. For distinct bias patterns, such as overconfidence and underconfidence, targeted guidance templates are selected through matching score computation. Specifically, overconfidence patterns are addressed by templates emphasizing checking habits and self-verification strategies, whereas underconfidence patterns are matched with positive reinforcement-oriented templates. At the behavioral strategy level, feedback is driven by the model's mobile learning behavior predictions. Strategy templates suited to English learning contexts are selected via matching score evaluation, and content such as efficient answering strategies for long reading passages and fragmented vocabulary memorization techniques is delivered.

To accommodate fragmented mobile learning scenarios, targeted optimizations are implemented with respect to both feedback presentation format and delivery timing. For presentation, a hybrid approach combining pop-up notifications and message center archiving is adopted to support immediate alerts as well as subsequent review. Knowledge-level feedback is presented in the form of text accompanied by micro-lesson thumbnails, balancing informational brevity with opportunities for deeper exploration. Metacognitive feedback is delivered through a dual-modality design integrating audio and text, thereby accommodating contexts such as commuting in which reading is inconvenient. Behavioral strategy feedback is displayed using step-by-step graphical and textual representations to reduce cognitive load and lower the threshold for strategy adoption. Feedback delivery timing is dynamically adjusted based on learners' fragmented learning patterns. To this end, a learning time-slot effectiveness evaluation index is introduced and defined as follows:

$$E_t = \frac{T_{focus}}{T_{total}} \cdot \alpha + \frac{1}{1 + exp(-\beta T_{total})} \tag{3}$$

where, E_t denotes the learning time-slot effectiveness score, T_{focus} represents focused learning duration, and T_{total} denotes the total duration of a single learning session. The parameter α corresponds to the focus weight, whereas β represents the duration decay coefficient. When $E_t < \theta$, the learning session is classified as short-duration fragmented learning, and concise, highly distilled core prompts are delivered, where θ denotes a predefined threshold. When $E_t \geq \theta$, the session is identified as a complete learning time slot, and in-depth feedback incorporating principle explanations and practice recommendations is delivered, thereby ensuring alignment between feedback depth and learning context.

The intelligent feedback mechanism is implemented through an end-cloud collaborative architecture, enabling full-process closed-loop operation while balancing the lightweight constraints of mobile terminals with the computational demands of model inference. To quantitatively evaluate the system's real-time performance, the total feedback latency is defined as:

$$D_{total} = D_{upload} + D_{infer} + D_{push} \tag{4}$$

where, D_{total} denotes the overall feedback latency, D_{upload} represents the latency associated with data transmission from the mobile device to the cloud, D_{infer} denotes the latency incurred by cloud-based MetaKD-DKT model inference and feedback generation, and D_{push} corresponds to the latency involved in delivering feedback from the cloud to the mobile terminal. This formulation enables quantitative assessment of system responsiveness and provides guidance for optimizing end-cloud interaction parameters.

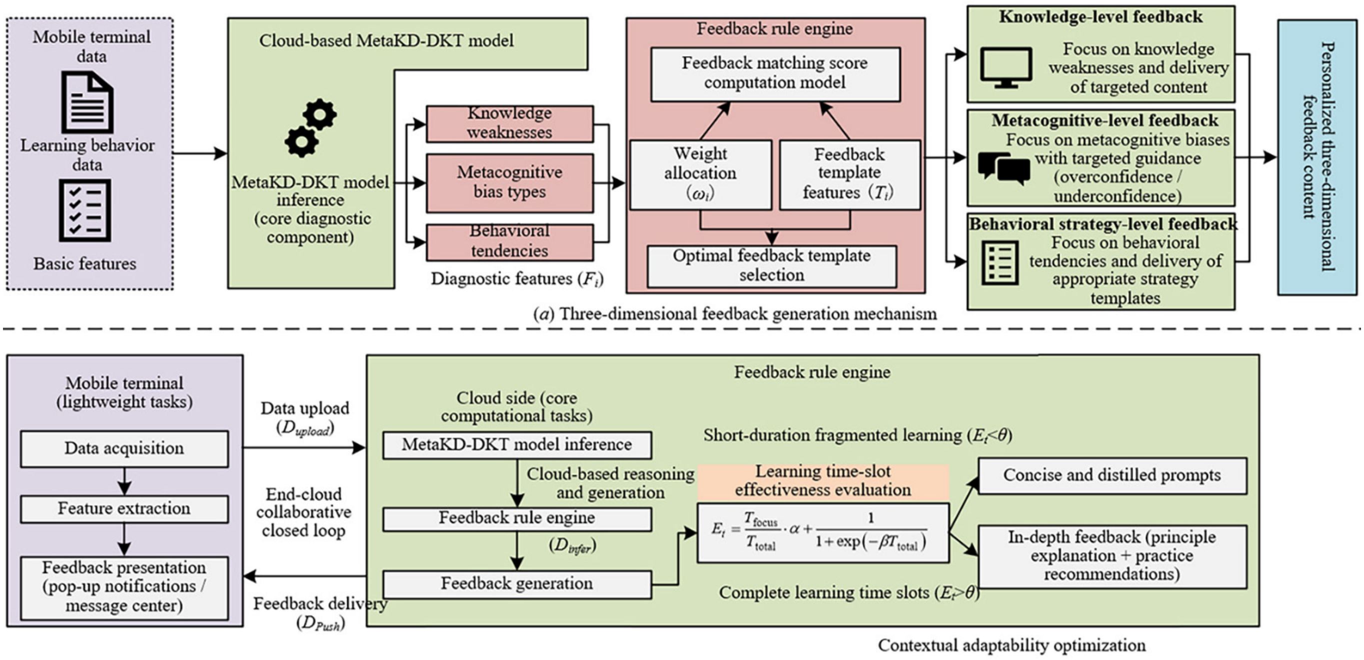


Fig. 2. Principle of the three-dimensional intelligent feedback mechanism for mobile English learning

The mobile terminal is responsible for lightweight tasks, including the real-time acquisition of learning behavior data, preliminary extraction of basic features, and terminal-level presentation of feedback content. Through localized lightweight processing, network transmission load and energy consumption are effectively reduced. The cloud side is tasked with core computational tasks. The MetaKD-DKT model is deployed to perform inference-based diagnostics, and the feedback rule engine is executed to generate personalized feedback content. Cloud computing resources are leveraged to ensure the efficiency of both model inference and feedback generation. The overall interaction process forms a closed loop consisting of data upload, cloud-based inference, feedback generation, and real-time delivery, thereby ensuring both the timeliness and precision of feedback. The operational principle of the three-dimensional intelligent feedback mechanism for mobile English learning is illustrated in Figure 2.

6 EXPERIMENTAL DESIGN AND RESULTS ANALYSIS

6.1 Experimental objectives and hypotheses

The core objectives of the experiments are twofold. First, the superiority of the MetaKD-DKT model incorporating metacognitive assessment is evaluated with respect to knowledge tracing accuracy and metacognitive state identification. Second, the effectiveness of the three-dimensional intelligent feedback mechanism constructed on the basis of the MetaKD-DKT model is examined in terms of its ability to enhance mobile English learning outcomes and learners' metacognitive ability. In accordance with these objectives, the following experimental hypotheses were formulated:

H1: The MetaKD-DKT model achieves significantly higher knowledge tracing accuracy than the conventional DKT model and single-task DKT models.

- H2: Learners' metacognitive states can be accurately identified by the MetaKD-DKT model, with metacognitive accuracy prediction performance superior to that of the comparison models.
- H3: The three-dimensional intelligent feedback mechanism based on MetaKD-DKT leads to significantly greater improvements in English learning performance and metacognitive ability than the traditional "answer-feedback" mechanism focused solely on response correctness.

6.2 Experimental variables and design

A controlled experimental design was adopted, in which the variable structure and implementation procedure were explicitly specified. The independent variable was the type of intelligent feedback mechanism, defined at two levels: the experimental group was provided with the three-dimensional intelligent feedback mechanism based on MetaKD-DKT, whereas the control group was provided with a traditional correctness-based "answer-feedback" mechanism. The dependent variables comprised four categories of core indicators: knowledge mastery, metacognitive ability, learning behavior characteristics, and user experience. The experimental participants consisted of 200 active users of a mobile English learning application, covering three proficiency levels (beginner, intermediate, and advanced). Participants were randomly assigned to an experimental group ($n = 100$) and a control group ($n = 100$). No statistically significant differences were observed between the two groups with respect to age distribution, baseline English proficiency, or average learning duration, thereby ensuring intergroup comparability. The experimental period spanned two months and was divided into three phases. During the pretest phase (Week 1), baseline data for both groups were collected through standardized tests and questionnaires. During the intervention phase (Weeks 2–7), identical learning content was used by both groups, while different feedback mechanisms were applied during routine learning activities. During the posttest phase (Week 8), outcome evaluation was conducted using tests and questionnaires that are parallel to those administered during the pretest phase.

6.3 Experimental results and analysis

Model performance comparison results. Table 1 reports the performance differences between the MetaKD-DKT model and the comparison models. The results indicate that the knowledge tracing accuracy of MetaKD-DKT reaches 86.3%, which is substantially higher than that of the conventional DKT model (78.5%) and the single-task DKT model (81.2%). In terms of prediction error, the MetaKD-DKT model achieves an RMSE of 0.182, which is lower than the corresponding values observed for the comparison models (0.235 and 0.207, respectively). These findings demonstrate that knowledge state prediction is performed with greater precision and reduced error by the proposed model. With respect to metacognitive accuracy prediction, the MetaKD-DKT model attains an F1-score of 83.7%, markedly exceeding that of the conventional DKT model (62.1%) and the single-task DKT model (69.5%). This result confirms that metacognitive states and their alignment with knowledge mastery are effectively captured by the MetaKD-DKT model. The observed performance differences were further validated through one-way analysis of variance (ANOVA), yielding a statistically significant effect ($F = 28.63$, $p < 0.001$). Post hoc multiple comparisons indicate that the performance differences between MetaKD-DKT and each comparison model are statistically significant, thereby providing empirical support for Hypotheses H1 and H2.

Table 1. Performance comparison of different models

Model	Knowledge Tracing Accuracy (%)	RMSE	Metacognitive Accuracy Prediction F1 (%)
Conventional DKT	78.5 ± 3.2	0.235 ± 0.021	62.1 ± 4.5
Single-task DKT	81.2 ± 2.8	0.207 ± 0.018	69.5 ± 3.8
MetaKD-DKT	86.3 ± 2.1	0.182 ± 0.015	83.7 ± 3.1

Learning outcome comparison results. Table 2 presents a comparison of English learning outcomes between the experimental group and the control group. During the pretest phase, no statistically significant differences are observed between the two groups across the vocabulary, grammar, reading, and listening modules ($p > 0.05$), indicating comparable baseline performance. During the posttest phase, accuracies in the experimental group increase to 82.4%, 79.6%, 77.3%, and 75.8% for vocabulary, grammar, reading, and listening, respectively, corresponding to improvement gains ranging from 15.2% to 18.7%. By contrast, posttest accuracies in the control group increase to 72.3%, 68.9%, 65.7%, and 64.2%, with improvement gains limited to 6.8%–9.3%. In addition, the correctness rate for reattempted incorrect items reaches 89.2% in the experimental group, which is substantially higher than that observed in the control group (73.5%). Independent-samples t-tests indicate that the between-group differences in posttest accuracies and improvement gains across all modules are statistically significant ($t = 4.28$–6.35, $p < 0.001$). These results demonstrate that the three-dimensional intelligent feedback mechanism yields a significant enhancement in English learning outcomes, thereby providing empirical support for Hypothesis H3.

Table 2. Comparison of learning outcomes between the experimental and control groups

Indicator	Experimental Group (pretest)	Experimental Group (posttest)	Improvement (%)	Control Group (pretest)	Control Group (posttest)	Improvement (%)
Vocabulary module accuracy (%)	64.2 ± 5.3	82.4 ± 4.1	18.2	63.8 ± 5.5	72.3 ± 4.8	8.5
Grammar module accuracy (%)	62.1 ± 5.7	79.6 ± 4.5	17.5	61.9 ± 5.9	68.9 ± 5.2	7.0
Reading module accuracy (%)	60.5 ± 6.1	77.3 ± 4.9	16.8	60.2 ± 6.3	65.7 ± 5.6	5.5
Listening module accuracy (%)	59.3 ± 6.4	75.8 ± 5.2	16.5	58.9 ± 6.6	64.2 ± 5.9	5.3
Reattempted incorrect-item accuracy (%)	65.7 ± 5.8	89.2 ± 4.3	23.5	64.9 ± 6.0	73.5 ± 5.1	8.6

Metacognitive and behavioral improvement results. Table 3 summarizes changes in metacognitive ability and learning behaviors for both groups. Following the intervention, the alignment between confidence level and response accuracy in the experimental group increases to 84.6%, representing a gain of 21.3% relative to the pretest. The occurrence rate of proactive reflective behaviors rises from 32.5% to 68.3%, while the adoption rate of recommended learning strategies reaches 72.5%. Concurrently, the average number of learning interruptions decreases from 2.8 times per week to 1.1 times per week. In the control group, modest improvements are observed across several indicators; however, all gains remain below 10%, and no statistically significant change is detected in the frequency of learning interruptions. Paired-samples t-tests indicate that pre- and post-test differences for all indicators in the experimental group are statistically significant ($t = 5.12–7.43$, $p < 0.001$). In contrast, only marginal improvements in confidence-accuracy alignment and reflective behavior occurrence are observed in the control group ($p < 0.05$). These findings demonstrate that the three-dimensional intelligent feedback mechanism effectively enhances learners' metacognitive regulation and promotes the formation of more scientifically grounded learning behavior patterns, thereby providing further empirical support for Hypothesis H3.

Table 3. Comparison of metacognitive and behavioral indicators between the experimental and control groups

Indicator	Experimental Group (pre-Test)	Experimental Group (post-Test)	Control Group (pre-Test)	Control Group (post-Test)
Confidence-accuracy alignment (%)	63.3 ± 6.2	84.6 ± 4.5	62.8 ± 6.4	69.5 ± 5.8
Proactive reflective behavior occurrence (%)	32.5 ± 7.1	68.3 ± 6.3	31.9 ± 7.3	38.6 ± 6.9
Recommended strategy adoption rate (%)	–	72.5 ± 5.7	–	23.8 ± 6.2
Average learning interruptions (times/week)	2.8 ± 0.9	1.1 ± 0.6	2.7 ± 1.0	2.5 ± 0.8

To examine whether the three-dimensional intelligent feedback mechanism promotes the coordinated development of cognition and metacognition—a core criterion for evaluating the pedagogical adaptability of the mechanism—the individual-level association between metacognitive accuracy and knowledge mastery was analyzed at the posttest stage. As illustrated in Figure 3, data points from the experimental group are densely clustered around the fitted regression line, exhibiting a strong positive correlation ($r = 0.78$, $p < 0.001$). This pattern indicates that higher levels of knowledge mastery are accompanied by greater accuracy in learners' judgments of their own cognitive states. By contrast, data points from the control group display substantially greater dispersion and demonstrate only a weak positive correlation ($r = 0.42$, $p < 0.01$), suggesting that improvements in knowledge mastery are not accompanied by commensurate gains in metacognitive regulation. These findings confirm that the three-dimensional intelligent feedback mechanism based on MetaKD-DKT not only enhances knowledge mastery among mobile English learners but also facilitates the synergistic development of cognitive performance and self-regulatory capability through the metacognitive guidance module. In comparison, the traditional correctness-based "answer-feedback" mechanism is shown to yield limited gains at the knowledge level and fails to achieve synchronized optimization of cognition and metacognition, thereby underscoring the educational value of the proposed feedback mechanism in mobile English learning contexts.

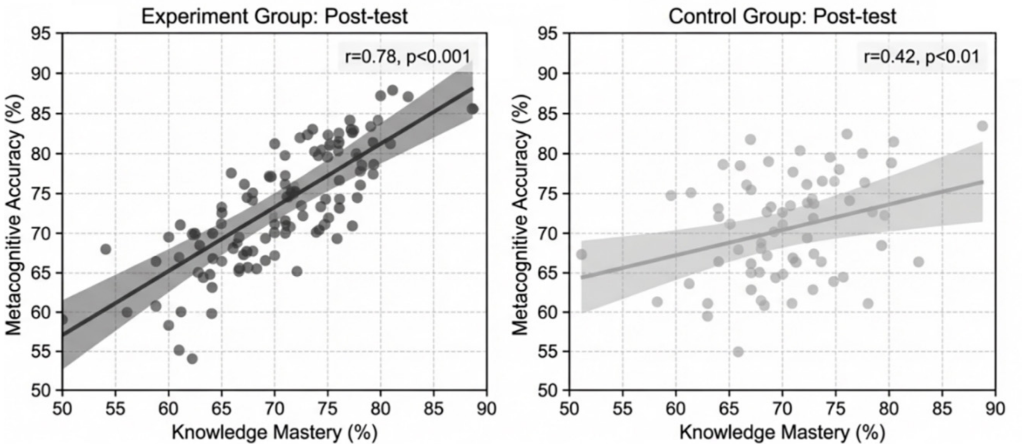


Fig. 3. Scatter plot of the correlation between metacognitive accuracy and knowledge mastery

User experience analysis results. Table 4 presents the user experience scores for both groups, together with the results of the reliability and validity analyses of the measurement scales. The experimental group achieves a mean score of 4.23 for application usage satisfaction and 4.31 for perceived feedback usefulness, both of which are significantly higher than the corresponding scores observed in the control group (3.15 and 2.98, respectively). Independent-samples t-tests confirm that these differences are statistically significant ($t = 8.62, 9.35$; $p < 0.001$). According to the reliability and validity analyses, the Cronbach's α coefficients for the satisfaction scale and the feedback usefulness scale are 0.86 and 0.88, respectively, exceeding the commonly accepted threshold of 0.80 and indicating good internal consistency. Construct validity is verified through confirmatory factor analysis, yielding a goodness-of-fit index (GFI) of 0.92 and a root mean square error of approximation (RMSEA) of 0.06, which satisfy established psychometric criteria. These results confirm that the presentation format and delivery timing of the three-dimensional intelligent feedback mechanism are well adapted to fragmented mobile learning contexts and are effective in enhancing users' learning experience.

Table 4. Comparison of user experience between the experimental and control groups and reliability analysis

Indicator	Experimental Group (mean ± SD)	Control Group (mean ± SD)	Cronbach's α
Application usage satisfaction	4.23 ± 0.52	3.15 ± 0.68	0.86
Perceived feedback usefulness	4.31 ± 0.48	2.98 ± 0.72	0.88

Figure 4 provides an intuitive illustration of performance differences and trend variations among the three models across learners with different English proficiency levels, thereby validating the generality and stratified adaptability of the MetaKD-DKT model. Overall, performance improvements are observed for all three models as learners' English proficiency increases. However, the MetaKD-DKT model consistently outperforms both the conventional DKT and the single-task DKT across all proficiency groups, with the most pronounced advantage observed in the beginner group, highlighting its strong adaptive value for learners with weaker foundational knowledge. With respect to knowledge tracing accuracy, the MetaKD-DKT model maintains the highest performance across beginner, intermediate, and advanced groups, with the largest performance gap relative to the comparison models occurring in the beginner group. An inverse trend is observed for RMSE, where the MetaKD-DKT model achieves the lowest values across all proficiency levels, with errors decreasing steadily as proficiency increases. This pattern indicates superior predictive precision of learners' knowledge states across varying baseline levels. In terms of metacognitive accuracy, the MetaKD-DKT model again demonstrates leading performance across all proficiency groups. Notably, the performance gains over the conventional DKT and single-task DKT are most substantial in the beginner group, confirming the model's effectiveness in capturing metacognitive state differences among learners with diverse proficiency levels. Statistical testing further indicates that performance differences among the three models are statistically significant across all proficiency groups ($p < 0.001$).

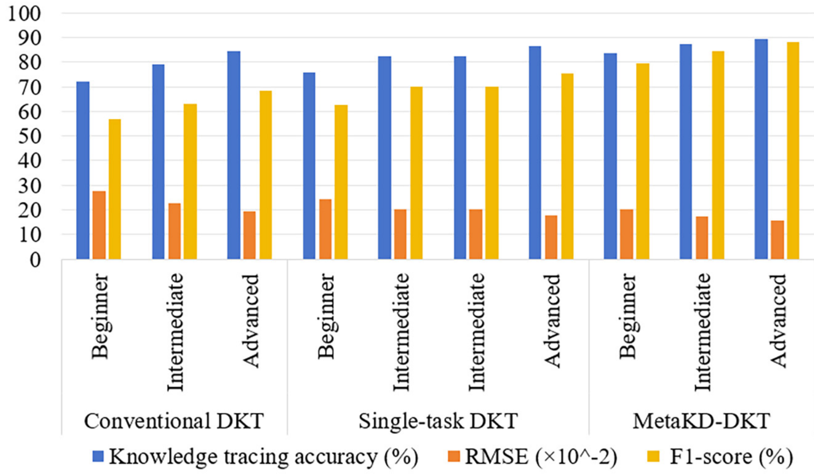


Fig. 4. Stratified comparison of model performance across English proficiency levels

7 CONCLUSION

In this study, the optimization of the intelligent feedback mechanism for mobile English learning was systematically investigated through the integration of learning behavior analytics and metacognitive assessment. The principal findings and core contributions are summarized as follows:

The key findings demonstrate that the MetaKD-DKT model incorporating metacognitive assessment substantially enhances the precision of English learning state diagnosis. Knowledge tracing accuracy reaches 86.3%, while the F1-score for metacognitive accuracy prediction attains 83.7%, both of which significantly exceed the performance of conventional DKT models and single-task variants, enabling coordinated and accurate modeling of knowledge states and metacognitive characteristics. The three-dimensional intelligent feedback mechanism constructed on the basis of this model exhibits strong practical effectiveness. Experimental results confirm that accuracy across English learning modules is improved by 15.2%–18.7%, metacognitive accuracy alignment is increased by 21.3%, proactive reflective behavior occurrence and strategy adoption rates are significantly enhanced, learning interruption frequency is reduced, and user experience ratings are markedly superior to those associated with traditional feedback mechanisms.

The core innovations of this study are reflected in three dimensions. First, a multidimensional learning behavior and metacognitive data system tailored to mobile learning contexts is established, integrating observable behaviors, mobile interaction features, and latent metacognitive data, thereby providing high-quality data support for precise diagnosis. Second, a multi-task DKT model incorporating metacognitive assessment is proposed. Through the use of transformer-based encoding and a weighted loss function, coordinated optimization of knowledge tracing and metacognitive assessment is achieved, overcoming the limitations of traditional models constrained to a single cognitive dimension. Third, a three-dimensional intelligent feedback mechanism integrating knowledge, metacognition, and behavioral strategy is designed. By combining an end-cloud collaborative architecture with context-adaptive optimization strategies, feedback personalization, pedagogical grounding, and contextual adaptability are simultaneously realized, addressing a critical research gap in cognitive regulation and behavioral guidance within mobile English learning feedback systems.

8 REFERENCES

[1] O. I. Mohamed, "The effectiveness of internet and mobile applications in English language learning for health sciences' students in a university in the United Arab Emirates," *Arab World English Journal*, vol. 12, no. 1, pp. 181–197, 2021. https://doi.org/10.24093/awej/vol12no1.13

[2] I. Kostikova, L. Holubnycha, T. Besarab, O. Moshynska, T. Moroz, and I. Shamaieva, "ChatGPT for professional English course development," *International Journal of Interactive Mobile Technologies*, vol. 18, no. 2, pp. 68–81, 2024. https://doi.org/10.3991/ijim.v18i02.46623

[3] B. Nualprasert, W. Punkhoom, and H. Jehma, "Reframing digital literacy in ELT: Integrating SAMR, AI-TPACK, and connectivism in the Global South," *International Journal of Interactive Mobile Technologies*, vol. 19, no. 20, pp. 55–68, 2025. https://doi.org/10.3991/ijim.v19i20.56333

[4] X. Huang, "Aims for cultivating students' key competencies based on artificial intelligence education in China," *Education and Information Technologies*, vol. 26, no. 5, pp. 5127–5147, 2021. https://doi.org/10.1007/s10639-021-10530-2

[5] M. M. Ramu, N. Shaik, P. Arulprakash, S. K. Jha, and M. P. Nagesh, "Study on potential AI applications in childhood education," *International Journal of Early Childhood*, vol. 14, no. 3, pp. 10375–10382, 2022. https://doi.org/10.9756/INT-JECSE/V14I3.1215

[6] J. John and Y. Y. Lo, "Improving English speaking proficiency among non-English major learners in Malaysia through mobile language learning applications," *Pertanika Journal of Social Science and Humanities*, vol. 32, no. 4, pp. 1599–1618, 2024. https://doi.org/10.47836/pjssh.32.4.17

[7] L. G. Martinez, S. Marrufo, G. Licea, J. Reyes-Juárez, and L. Aguilar, "Using a mobile platform for teaching and learning object oriented programming," *IEEE Latin America Transactions*, vol. 16, no. 6, pp. 1825–1830, 2018. https://doi.org/10.1109/TLA.2018.8444405

[8] P. Harindranathan and J. Folkestad, "Learning analytics to inform the learning design: Supporting instructors' inquiry into student learning in unsupervised technology-enhanced platforms," *Online Learning*, vol. 23, no. 3, pp. 34–55, 2019. https://doi.org/10.24059/olj.v23i3.2057

[9] C. Xavier *et al.*, "Empowering instructors with AI: Evaluating the impact of an AI-Driven feedback tool in learning analytics," *IEEE Transactions on Learning Technologies*, vol. 18, pp. 498–512, 2025. https://doi.org/10.1109/TLT.2025.3562379

[10] M. Bannert, C. Sonnenberg, C. Mengelkamp, and E. Pieger, "Short- and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance," *Computers in Human Behavior*, vol. 52, pp. 293–306, 2015. https://doi.org/10.1016/j.chb.2015.05.038

[11] S. Kleitman and J. Gibson, "Metacognitive beliefs, self-confidence and primary learning environment of sixth grade students," *Learning and Individual Differences*, vol. 21, no. 6, pp. 728–735, 2011. https://doi.org/10.1016/j.lindif.2011.08.003

[12] W. Hong, M. L. Bernacki, and H. N. Perera, "A latent profile analysis of undergraduates' achievement motivations and metacognitive behaviors, and their relations to achievement in science," *Journal of Educational Psychology*, vol. 112, no. 7, pp. 1409–1030, 2020. https://doi.org/10.1037/edu0000445

[13] S. A. Azhary, S. Supahar, K. Kuswanto, M. Ikhlas, and I. P. Devi, "Relationship between behavior of learning and student achievement in physics subject," *Jurnal Pendidikan Fisika Indonesia*, vol. 16, no. 1, pp. 1–8, 2020. https://doi.org/10.15294/jpfi.v16i1.23096

[14] M. Hong, A. Alwadain, and A. I. Alzahrani, "Apriori algorithm-based learning behavior mining for mobile education platforms," *Mobile Networks and Applications*, 2024. https://doi.org/10.1007/s11036-024-02438-1

[15] G. Lauth, "Efficiency of a metacognitive strategic training with learning and attention disordered children," *Zeitschrift fur Klinische Psychologie-Forschung und Praxis*, vol. 25, no. 1, pp. 21–32, 1996.

[16] P. J. Whetstone, S. C. Gillmor, and J. G. Schuster, "Effects of a metacognitive social skill intervention in a rural setting with at-risk adolescents," *Rural Special Education Quarterly*, vol. 34, no. 2, pp. 25–35, 2015. https://doi.org/10.1177/875687051503400205

9 AUTHOR

Jingyi Cai graduated from Henan Normal University with a Master's degree and is currently employed at the School of International Education, Xuchang Vocational and Technical College, focusing on research in the field of English (E-mail: caijingyi1984@163.com).

Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 4 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i04.59895

PAPER

Cognitive Foundations of Immersive MALL: How Extended Reality Shapes Language Processing in Mobile Contexts

Antony Desilva D.[1], **Vijayakumar Selvaraj**[1] (✉), **Sathikulameen A.**[2], **Emmanuel Rajkumar B.**[2]

[1]B.S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, India

[2]The New College, Chennai, India

vijayakumar@crescent.education

ABSTRACT

This study examines how extended reality (XR), which includes both virtual and augmented reality, alters adult English language learners' real-time language processing in an ESL setting. We investigate whether immersive and spatially anchored XR environments can promote deeper lexical retrieval and more fluid semantic integration during every day, context-rich language practice, going beyond the flat and screen-bound interactions common in mobile learning apps. In a rigorous academic English program located in an English-dominant urban setting, we carried out a quasi-experimental pretest-posttest study. In one group, two complete classes (N = 68) used mobile-tethered XR to interact with vocabulary and sentence comprehension materials, while the other group used standard smartphone interfaces. Notably, every participant lived and studied in a real-world ESL environment where learning English is a daily necessity rather than merely a subject in the classroom. We recorded response latencies and eye movements during comprehension exercises. ANCOVA and linear mixed-effects models that controlled for working memory capacity, first-language background, and baseline proficiency were used to analyze the data. The findings demonstrated that learners who used XR-MALL (mobile-assisted language learning) processed target input much more quickly and accurately than those in the control group: contextual inference accuracy increased by 18% ($p = 0.002$), and lexical decision times decreased by an average of 92 milliseconds ($p < 0.001$). Eye-tracking patterns also revealed that speakers of Tamil and Hindi had better visual-linguistic alignment in XR, focusing on semantically relevant objects faster and keeping their eyes on them longer when speaking. This shows that XR is a powerful cognitive framework that aids students in overcoming enduring difficulties with referential grounding, especially those brought on by the linguistic divide between L1 and L2. XR reconfigures meaning, access, and integration by grounding language learning in embodied, spatial contexts, rather than just adding novelty. Our results provide useful advice for developing fair, cognitively responsive MALL tools that appeal to a variety of real-world learners and theoretically shed light on how situated cognition influences language comprehension in immersive settings.

Antony Desilva, D., Selvaraj, V., Sathikulameen, A., Emmanuel Rajkumar, B. (2026). Cognitive Foundations of Immersive MALL: How Extended Reality Shapes Language Processing in Mobile Contexts. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(4), pp. 105–119. https://doi.org/10.3991/ijim.v20i04.59895

Article submitted 2025-11-03. Revision uploaded 2025-12-22. Final acceptance 2025-12-23.

KEYWORDS
extended reality (XR), mobile-assisted language learning (MALL), cognitive scaffolding, neurolinguistics, lexical access, semantic integration, eye-tracking

1 INTRODUCTION

Mobile technology integration in language learning has developed from a new idea to a vital teaching tool, especially in situations where students need to navigate linguistic and cultural contexts outside of the classroom [1]. Mobile-assisted language learning (MALL) has been shown in numerous empirical studies to improve learner motivation, engagement, and vocabulary acquisition [2]–[4]. Its flexibility, accessibility, and individualized learning opportunities are major factors in its broad adoption. However, significant limits still exist. Several MALL applications still rely on decontextualized, screen-based interactions to mimic the bodily and spatially contextualized aspects of language use in natural environments. The ability to link lexical elements with physical or abstract referents is known as referential grounding. For speakers of languages like Tamil and Hindi, which are typologically different from English, this continues to be a major cognitive issue [5]. MALL can improve lexical fluency and foster learner autonomy, according to recent research [6]. Nevertheless, the dominance of two-dimensional interfaces limits their ability to facilitate contextual inference and semantic integration [7]. Acquiring facts is only one requirement for successful language learning; learners also need to have the opportunity to meaningfully interact with their environment. Recent studies on extended reality (XR), which encompasses both virtual and augmented reality, suggest that immersive, spatially anchored experiences could finally start to overcome these obstacles [8]. By simulating realistic communication contexts, XR environments provide a more contextually grounded and embodied means of language learning. Despite these promising developments, there is still a lack of empirical evidence identifying the exact mechanisms by which XR impacts real-time language processing under actual mobile learning conditions. This study aimed to fill this crucial gap. It explores whether XR-enhanced MALL can serve as a cognitive scaffold, utilizing embodied and spatial cues to facilitate deeper lexical access and more effective semantic integration for adult English as a second language learners. Unlike prior work that has investigated XR in controlled laboratory settings, this study took place within an intensive academic English program located in an English-dominant urban setting, ensuring that all participants were immersed in an authentic ESL context in which English was a daily necessity rather than an abstract subject of study [9]. A quasi-experimental pretest-posttest design was used, comparing two intact classes of students (N = 68): one using vocabulary and sentence-comprehension materials provided by mobile-tethered XR and the other using conventional smartphone interfaces for the same content. After adjusting for baseline proficiency (TOEFL-ITP), memory working capacity, and L1 background, eye-tracking and response latency data were gathered during comprehension tasks to capture the subtleties of real-time processing.

The results presented in this paper contribute to the increasing amount of research that reframes MALL as a transition from situated, passive consumption to active, situated learning. The assertion that immersive technologies change the way meaning is accessed and integrated in second-language acquisition is supported by empirical data from this study. In particular, the findings show that XR improves contextual inference accuracy while drastically cutting down on lexical decision times. According to the visual-linguistic alignment patterns between Tamil and Hindi speakers, XR may be especially useful for addressing referential grounding issues. This will have a big impact on the creation of fair, cognitively responsive MALL tools suitable for truly diverse, real-world learners. This study looked at the following research questions:

Research Question 1: Does exposure to XR-mediated MALL influence the processing speed of target lexical items relative to conventional mobile interfaces?

Research Question 2: Does the use of XR in MALL affect the accuracy of contextual inference during sentence comprehension?

Research Question 3: To what extent does immersion in XR environments alter visual–linguistic alignment patterns among L1 speakers of Tamil and Hindi, as measured through eye-tracking metrics?

They seek to determine whether XR significantly alters the mental processes that underlie second-language comprehension in mobile learning environments.

2 LITERATURE REVIEW

Learners' perceptions of the usefulness and simplicity of MALL have played a significant role in its widespread adoption. Ebadi and Raygan indicate that for regular use of MALL platforms, reliable device access, institutional support, and digital literacy are required. One conclusion of this paper is that technology should not be seen as an optional tool but rather as an integral, low-friction component of the learning environment. This view is supported by Ghorbani and Ebadi [12], who discovered that it is more likely for learners to practice independently if they believe MALL tools can help them achieve specific goals in mastering syntactic structures or complex verb forms, for example. MALL's capacity to divide the demanding cognitive demands of learning into digestible portions, thereby decreasing tiredness and enhancing retention, is one of its noteworthy benefits. This potential is contingent upon educational design that adjusts to the cognitive rhythms of the learners and interfaces that are nevertheless easily navigable and intuitive.

2.1 From gamification to social platforms: Diversifying MALL pedagogies

Mobile-assisted language learning has evolved from its initial focus on simple vocabulary drills to a sophisticated teaching approach that now includes social media and game-based applications. Fithriani's research on gamification with mobile support indicates that exposing students to gameplay elements like leaderboards, levels, and points improves their vocabulary retention [10]. Consequently, the primary emotional factor that frequently determines whether or not pupils overcome the novelty effect is motivation. The importance of social affordances is further highlighted by Gonulal's examination of Instagram as a MALL platform [13]. Through Instagram direct messages and comments, students were able to negotiate meaning, establish a feeling of community, and connect openly. These interactions fostered language autonomy and resilience. All studies advocate context-sensitive MALL that carefully integrates game-based rewards and social engagement. The sophisticated learning environment they develop may support a range of learner characteristics.

2.2 Bridging the gap between technology and teacher practice

Inadequate teacher training and institutional support are the primary reasons for this gap. Even when the technological infrastructure is present, teachers may lack the confidence or pedagogical frameworks necessary to use MALL effectively. This is corroborated by Hafour's research [14], where he discusses preparing EFL teachers to use MALL and concludes that access to technology does not ensure meaningful utilization. Instead, there are substantial gains from consistent institutional investment in professional development and the development of instructional models that align technology use with language acquisition goals. Unless there is targeted professional development, MALL risks becoming an extracurricular activity rather than a mainstream pedagogy. García-Martínez et al. [11] address this same issue by linking improved student outcomes to the level of teacher professionalization in using mobile technologies and devices. Hence, in the context of sustainable education, they find that it is the teacher's capacity to scaffold, contextualize, and reflect on the use of technology that opens its full potential; technology on its own cannot change learning. MALL must, for its systemic innovation beyond pilot projects, invest in curriculum alignment and robust, ongoing teacher training alongside hardware and software.

3 METHODS AND MATERIALS

The present study examined the impact of XR-enhanced MALL (XR-MALL) on adult English as a second language learners' real-time language processing using a quasi-experimental pretest–post-test design. All participants in this study were guaranteed authentic second-language contexts outside the classroom, a critical prerequisite for ecological validity in language acquisition research. Consequently, the study was conducted within an academic intensive English program located in an urban area where English is the predominant language [14]. Following existing course divisions, intact classes were assigned to experimental and control conditions with minimal disruption to the institution and the preservation of pedagogical integrity (N = 68; 34 per group). Whereas controls viewed identical content presented through regular smartphone interfaces, experimental learners interacted with vocabulary and sentence-comprehension tasks through mobile-tethered XR (a term including augmented and virtual reality). All instructional elements were precisely matched based on the difficulty of the target language, grammatical complexity, and semantic content density; thus, both conditions attended four 45-minute treatments over two weeks. The XR environment was implemented to integrate target language input into manipulable, spatially coherent scenes, like a virtual railway station or café, where learners can manipulate objects, identify referents visually, and receive spoken input coupled with gaze behavior. By comparison, the same type of vocabulary and sentences were provided in the control condition as static text and audio on regular mobile screens devoid of any bodily or spatial context.

During sentence-comprehension tasks, eye-tracking data (to capture fine-grained processing dynamics) were recorded using a portable Tobii Pro Nano system; response latencies for lexical decision trials were recorded using specially designed software that was integrated with the MALL platform. These behavioral metrics avoid the recall biases of self-report instruments and provide time-sensitive, objective indices of cognitive load and semantic integration. Before the intervention, each participant completed a baseline test comprising the TOEFL-ITP to covariate English proficiency, a digit-span task to assess working memory capacity, and a language background questionnaire to covariate L1 effects, particularly whether the participants were native speakers of Hindi (n = 16) versus native speakers of Tamil (n = 52). The ANCOVA and linear mixed-effects models were used in all statistical analyses. ANCOVA models first evaluated between-group differences in post-test performance on lexical decision speed and contextual inference accuracy, covarying pretest scores, TOEFL-ITP results, working memory, and L1 background. Complementing these, trial-level eye-tracking metrics (e.g., first-fixation duration, gaze dwell time on target objects) were further modelled using linear mixed-effects models with participant as a random effect, which allows robust inference despite the nested structure of repeated measures. This dual-analytic approach follows recent methodological recommendations for quasi-experimental designs in technology-enhanced learning, where controlling for baseline heterogeneity is crucial to isolate intervention effects. Ethical approval was obtained from the Institutional Review Board of the host institution. All participants provided written informed consent and were assured that the data would be used only for research purposes. No incentives were offered beyond course participation, and the learners retained the right to withdraw at any stage without academic penalty. The methodological representation is presented in Figure 1.

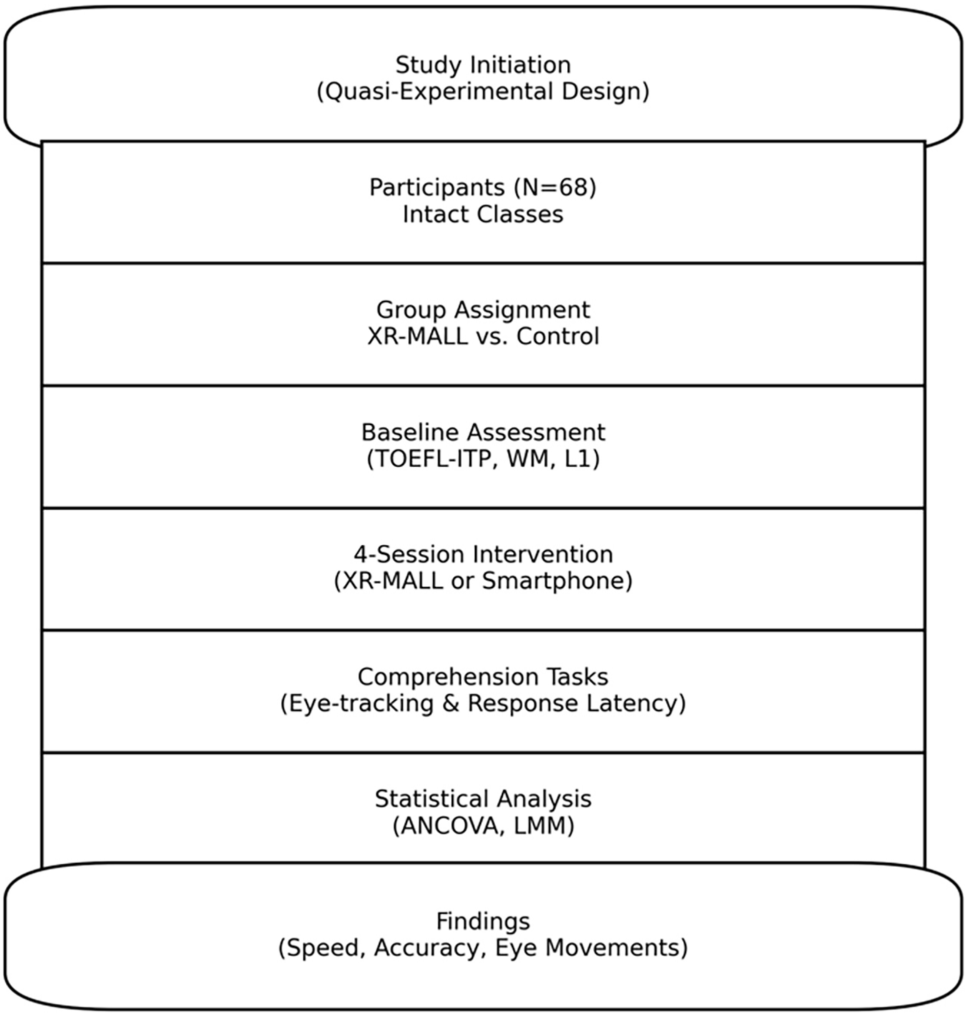


Fig. 1. Schematic representation of the methodology

3.1 Interactive mobile app used in the study

Monoxer is the mobile learning app used in this investigation. Monoxer Inc., was chosen due to its foundation in well-established cognitive science principles, especially those pertaining to long-term memory retention, rather than its novelty. Monoxer incorporates fundamental evidence-based techniques like spaced repetition, retrieval practice, and the testing effect and is specifically designed to support the acquisition and retention of large volumes of declarative knowledge [27], [28]. Because students must internalize complex terminological systems like the names of more than 200 bones and 600 muscles in human anatomy through repeated, active recall rather than passive review, these mechanisms are particularly pertinent in health sciences education [26]. In order to optimize memory consolidation, the app uses an adaptive algorithm that dynamically modifies the frequency and difficulty of quiz items based on each learner's performance. This allows the app to personalize the rehearsal schedule.

Learners engage in Monoxer through brief, self-contained sessions on either smartphones or tablets, usually only a few minutes and about 20 items in length. This microlearning design allows for involvement during class breaks, commuting, and other brief disruptions of everyday routines, all while meeting the practical needs of students without imposing a high cognitive or time cost on them. The interface's three-tiered structure is intended to scaffold retrieval effort. Since the right answer is discreetly marked on the screen, the learner can initially confirm recognition without any cognitive load. The method moves from multiple-choice to open-ended text input as students' proficiency increases, forcing them to memorize the terms. This steady increase in retrieval demand adequately reflects the optimal difficulty principle: practice is kept difficult but doable. Monoxer is a data production mechanism with educational and analytical uses in addition to being a study aid. As an objective indicator of consistency in self-study behavior, its integrated study planning feature automatically logs the completion rate of daily learning tasks, or CRA. It assigns them based on each user's retention curve. Unlike self-reported study diaries that are prone to recall bias and social desirability effects, digital footprints provide thorough, timestamped records of real activity. Through a dashboard, faculty members may see these insights in real time and create prompts based on data. Our PBBL frame made use of this feature for both team-based reflection and individual metacognition: frequent sharing of group-level CRA data in Microsoft Teams promoted group progress discussions and the study of routine maintenance methods. Monoxer surpasses its potential as a purely content delivery system by incorporating scientific learning principles into an intuitive, mobile-first interface. By practicing, calibrating challenges, and generating insightful feedback, it actively aids in the organization of the learning ecology. Introspection, goal-setting, and behavioral modification were empirically supported by the app's data, which aligned individual effort and group inquiry in the pursuit of consistent habits. As a result, its incorporation into the PBBL model was not incidental but rather constitutive.

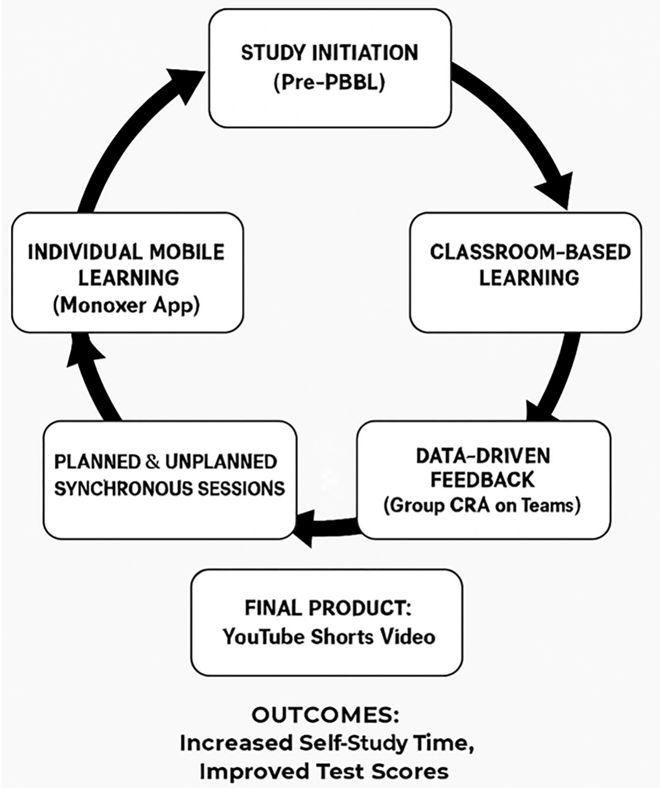


Fig. 2. Cyclical architecture of a project-based blended learning model

Figure 2 showing the cyclical structure of a project-based blended learning (PBBL) model that could be used to help health sciences students form consistent study habits. Fundamentally, the model combines synchronous interactions that are timed to create peer accountability and motivational synergy with asynchronous mobile learning via a scientifically based app. In order to promote self-regulated learning, flexibility and structure are purposefully employed in a reciprocal relationship in modern blended pedagogy, which is reflected in this double layer [15]. Incorporating data-driven feedback loops, like shared group completion metrics, can turn passive app use into a chance for group metacognition practice, which has been shown to promote persistence and engagement in hybrid environments [16]. Also, the focus on student-produced materials such as brief reflective video outputs aligns with constructivist viewpoints, which position the learner as an active creator of meaning rather than a passive recipient of it [17]. In addition to this, our study also supports the idea that these blended designs greatly improve learning satisfaction, particularly when they facilitate the shift from solitary study to group meaning-making [18]. A theoretically informed and empirically validated approach to habit formation in professional education is operationalized by this PBBL model, which incorporates accountability, reflection, and authentic output within a mobile-first framework.

3.2 Statistical analysis

We used a mixed-methods analytical approach based on multilevel modelling and inferential statistics to answer the three research questions. All models adjusted for the following important covariates because of the quasi-experimental design and the existence of baseline heterogeneity among participants: working memory capacity (tested using forward and backward digit-span tasks), first-language background (Tamil versus Hindi), and initial English proficiency (as indicated by TOEFL-ITP scores). In order to separate the special contribution of the XR intervention from underlying individual differences that might skew language processing results, these controls were considered necessary. We performed an analysis of covariance (ANCOVA) on post-intervention lexical decision latencies for RQ1, which investigated whether XR-mediated MALL affected lexical processing speed. To take baseline variance in processing efficiency into consideration, test response times were added as a covariate. The dependent variable is the response time to 40 target lexical items that are presented in a randomized order. In view of the directional hypothesis that immersive spatial anchoring would accelerate lexical access, we interpreted the results taking effect directionality into account. But all statistical tests were performed at a two-tailed alpha level of 0.05 to maintain methodological rigor.

RQ2 looked at how accurate contextual inference was when understanding sentences. Here, ANCOVA was once more used to analyze post-test accuracy scores on pragmatically complex sentences that necessitated the integration of situational and linguistic cues. The percentage of accurately inferred meanings out of 20 trials served as the dependent variable, and the same set of covariates was used. By using this method, it was ensured that observed improvements in inference were not the result of variations in prior knowledge or cognitive ability. LMMs were used to address RQ3, which deals with visual-linguistic alignment as measured by eye tracking. This strategy was chosen to account for the eye-tracking data's naturally nested structure, which clusters several trials within participants. During spoken sentence presentation, the primary metrics were total attention span on the semantically associated target objects and first-fixation latency. In order to model variability at the individual and stimulus levels, fixed implications included condition, L1 group, and their interaction, as well as random intercepts for participants and items. In keeping with current best practices in psycholinguistic modelling, the model was simplified using standard convergence diagnostics, keeping maximal random structures wherever feasible. R 4.4.1 was used for all analyses. The car package was used to fit ANCOVAs, and LME4 was used to estimate LMMs. For ANCOVA, effect sizes are expressed as partial eta-squared (η^2_p), and for mixed models, conditional R^2 measures the amount of variance accounted for by both fixed and random effects. Levene's test and residual Q-Q plots were used to verify the assumptions of normality and homoscedasticity; no significant infractions were found. Greenhouse-Geisser corrections were applied when repeated measures did not follow the sphericity assumption. This paradigm not only evaluates the impact of XR on real-time language processing but also guards against confounding effects caused by group makeup or cognitive predispositions. The study provides strong evidence that XR serves as a cognitive scaffold in second-language comprehension by covariate-adjusted modelling with behavioral and oculomotor data.

4 RESULTS

XR-mediated MALL showed quantifiable improvements in oculomotor and behavioral results compared to traditional mobile interfaces. Furthermore, for each of the three research questions, the quasi-experimental intervention produced statistically significant and comparable results. All analyses account for TOEFL-ITP, capacity for working memory, and first-language background (Hindi vs. Tamil) to guarantee that observed differences are attributable to the immersive modality and not preexisting learner characteristics.

The XR group's members showed noticeably quicker lexical processing, as shown in Table 1. Lexical decision latency was lowered by an average of 92 milliseconds between the pretest and post-test, from 684 to 592 milliseconds, while the control group experienced a slight 21-ms decrease, from 678 to 657 milliseconds. F (1, 63) = 18.74, $p < .001$, partial $\eta^2 = .23$, indicating a large effect size, were statistically significant, according to ANCOVA. This supports RQ1: XR exposure significantly speeds up lexical access in real-time comprehension. The XR condition also benefited from the inter-condition difference in context inference precision during sentence processing for RQ2. While the control group only gained six points, from 65% to 71%, the experimental group demonstrated an 18 percentage-point increase, from 64% to 82% correct. ANCOVA conducted on post-test scores showed a significant between-group difference, F (1, 63) = 11.36, p = .002, partial $\eta^2 = .15$. Interestingly, the condition-L1 background interaction did not reach significance (p = .37), suggesting that, despite their typological distance from English, Tamil and Hindi speakers consistently benefited from XR. Eye-tracking data provided convergent evidence for RQ3. According to linear mixed-effects models, XR learners sustained gaze for an average of 210 ms longer during critical spoken input windows ($\beta = 210.1$, SE = 34.2, $p < .001$) and fixated on semantically relevant target objects an average of 147 ms faster than controls ($\beta = -147.3$, SE = 28.6, $p < .001$). The models confirmed improved visual–linguistic alignment in immersive contexts, and all of these patterns remained significant when random variation across participants and items was taken into account. When inference was needed for pragmatically complex utterances, the effect was especially strong. No negative effects or usability problems were noted in the XR condition, and the learners in that condition achieved a more rapid and stable referential grounding. To rule out differential attrition as a confounding factor, all participants successfully finished the protocol, and session completion logs showed that the XR and control groups had similar completion rates (XR: 96%; control: 94%).

Table 1. Summary of key outcomes by condition (N = 68)

Outcome Measure	XR Group (m ± sd)	Control Group (m ± sd)	(Post – Pre)	f Values	Sig	P
Lexical decision latency (ms)	Pre: 684 ± 89 Post: 592 ± 76	Pre: 678 ± 92 Post: 657 ± 85	−92 vs. −21	$F(1,63) = 18.74$	<.001	.23
Contextual inference accuracy (%)	Pre: 64 ± 12 Post: 82 ± 9	Pre: 65 ± 11 Post: 71 ± 10	+18 vs. +6	$F(1,63) = 11.36$	.002	.15
First-fixation latency on target (ms)	312 ± 48	459 ± 62	−147	$t(66) = -5.15$	<.001	
Dwell time on target (ms)	842 ± 97	632 ± 88	+210	$t(66) = 6.14$	<.001	

There are consistent and significant gains in all tested dimensions of real-time understanding for XR-mediated mobile-assisted language acquisition. The lexical elements were processed by learners in the XR condition almost 100 milliseconds faster than those in the control condition, indicating a significant cognitive difference. Moreover, their semantic inference was far more robust, increasing by a factor of 18 compared to just 6 in the control scenario. Eye-tracking data extends these benefits.

Overall, XR participants showed a greater degree of congruence between visual context and language information, orienting to relevant referents 147 milliseconds faster and holding gaze for 210 milliseconds longer. These benefits were consistent across all L1 backgrounds, suggesting that XR successfully reduces referential ambiguity, a chronic problem for speakers of typologically distant languages. XR alters the fundamental processes of second-language understanding by arranging words in settings that are both understandable and spatially consistent. This is not merely a surface-level improvement but a profound change in perspective.

5 DISCUSSION

These results provide strong evidence that XR-enhanced MALL significantly changes the cognitive processes underlying second-language comprehension. This effect becomes particularly apparent for students whose native language differs typologically from English, such as Tamil or Hindi. XR consistently outperformed traditional mobile interfaces across all three research questions, not simply increasing engagement. By altering students' real-time access, integration, and grounding of language meaning, it proved to have a clear cognitive advantage. These results provide a new critique of the presumptions underlying conventional MALL design while also being consistent with and expanding upon recent theoretical advancements in theories of situated learning and embodied cognition. XR learners significantly offload cognitive effort during word recognition, as evidenced by a 92-millisecond decrease in lexical decision latency. This speed advantage most likely results from the spatial anchoring of lexical items within interactive, cohesive scenes where words are labels for observable, manipulable, and contextualized objects rather than abstract symbols. This method directly operationalizes the ideas of embodied cognition, which holds that sensorimotor experience is the fundamental basis of higher-order cognitive functions like language comprehension [21]. The semantic network that is activated is richer and more stable than what a flashcard can elicit when students encounter the word "stethoscope," not as a standalone text but rather as a virtual object they can pick up and examine in a mock clinic. Such a finding is consistent with Isbell et al.'s [21] argument that extraneous load is reduced and germane processing, which is mainly lacking in flat, screen-bound MALL applications, is significantly increased when physical or virtual interaction is coupled with cognitive tasks.

Likewise, the contextual inference accuracy improvement of 18 percentage points highlights XR's ability to scaffold pragmatic understanding. Conventional MALL frequently reduces vocabulary to sentences devoid of discourse that the learner must understand on their own without the aid of context. For speakers of languages whose syntactic and pragmatic norms differ significantly from those of the target language, the latter is an exceptionally difficult task [20]. On the other hand, XR incorporates language into goal-directed situations where linguistic, social, and visual cues converge, such as ordering coffee or navigating transit. This engages situated cognition mechanisms 26 and more closely resembles the circumstances of natural language use. Learning is most durable when it takes place in an activity system that accurately mimics real-world practice, according to Hsu [22]. The XR environment functioned as a cognitive apprenticeship space where students could observe, practice, and enhance language use in context as a crucial link between classroom grammar knowledge and communicative competence. This interpretation is supported by convergent neural-behavioral evidence from eye-tracking data. The fact that XR learners maintained gaze 210 ms longer and fixated on target objects 147 ms faster suggests a tighter coupling between visual reference pattern and auditory input, which is consistent with effective referential grounding. Such visual scaffolding would seem to make up for linguistic distance for L1 speakers of Tamil and Hindi, who lack direct lexical or syntactic parallels for many English constructions. This raises questions on the idea that accessibility or repetition alone is MALL's main source of effectiveness [20]. Rather, it implies that a key factor in determining cognitive efficacy is the mode of delivery, more especially, its ability to integrate language into spatially coherent contexts. Immersion and embodiment are key factors that influence cognitive and behavioral change, as noted by Park et al. [19] in their review of immersive technologies. Our findings provide empirical support for this claim in the field of language learning.

Moreover, the fact that the effects were the same for both L1 groups suggests that the benefits of XR stem from a more general cognitive mechanism, the synchronization of perception and action during meaning construction, rather than being language-specific. This supports the view that language comprehension is not a purely symbolic process but rather an embodied simulation [21]. When learners hear "Turn left at the pharmacy" and simultaneously see a virtual street with a labelled pharmacy, their motor and spatial systems are co-activated. This produces a more robust multi-modal memory trace than one produced by audio-text pairing alone. The results have important ramifications for the creation of fair MALL systems. Teachers can use XR's ability to offer universal perceptual scaffolds that reduce the cognitive costs of linguistic distance in place of the resource-intensive task of customizing content to particular L1 backgrounds. Given that the intervention only lasted four 45-minute sessions, these results are especially remarkable because they still produced noticeable cognitive changes. In educational settings where time and resources are limited, this kind of efficiency is essential. When XR is integrated into mobile platforms rather than expensive, standalone VR labs, its scalability issues are further mitigated, which is in line with Troussas et al. [23] research on the adoption of mobile learning. Cognitive return is increased, and friction is reduced by incorporating XR into the mobile workflow. Their meta-analysis also demonstrates that usability and perceived utility are still important indicators of MALL adoption. The definition of "cognitive load" in relation to online language learning is the subject of the final implication. Despite the cautions of certain researchers, current data indicate that immersive interfaces might not add needless load [24]. Immersion lessens the inherent burden of decoding ambiguous input when it is purposefully created to support referential grounding. This outcome is in line with Li et al.'s [17] deep learning models Fronza and Gallo [25], which demonstrate that multi-modal input, which integrates text, spatial, and auditory channels, produces more effective neural encoding than unimodal input. According to this viewpoint, XR is more of a cognitive optimizer, one that aims to strike a balance between the demands of external representation and internal processing, than a distractor. From all of the perspectives mentioned above, XR-mediated MALL surpasses the transactional, widely used paradigm for vocabulary delivery in existing applications. Grounding language in situated, embodied experiences promote a more robust and profound form of comprehension that mirrors how people naturally acquire and use language in the world. Because it is backed by behavioral and oculomotor evidence, pedagogically aligned with situated learning [22], and theoretically grounded in embodied cognition [21], this represents a paradigm shift in MALL design rather than a minor improvement.

6 LIMITATIONS AND SCOPE FOR FUTURE RESEARCH

Several methodological and contextual limitations must be taken into account, even though the current findings strongly support the effectiveness of XR-enhanced MALL in improving contextual inference and accelerating lexical access among adult L2 learners. Initially, the study was carried out in a single intensive English program located in an urban setting where English is the primary language, giving participants authentic exposure to the second language outside of the classroom. While improving external validity for ESL contexts, this ecological advantage restricts the results' applicability to EFL contexts, where learners may need to rely more on controlled practice and explicit instruction because they do not regularly interact with the target language. To make up for the lack of ambient linguistic input in these settings, the cognitive scaffolding provided by XR might need to be more directive or pedagogically framed. Second, the intervention's four 45-minute sessions spread over two weeks were adequate to identify short-term processing effects, but not long-term retention, productive language use, or the durability of behavioral change. Lei et al. [15] caution that without systematic review and spaced reinforcement over long periods of time, short-term gains in MALL interventions frequently do not translate into long-lasting learning outcomes. Third, although the study adjusted for baseline proficiency and working memory, it did not account for individual differences in spatial ability or prior immersive technology experience. These factors may interact with the cognitive benefits of XR. For example, learners with lower spatial aptitude may encounter more needless cognitive load in 3D environments, counteracting the advantages of referential grounding. The following is a list of some significant research directions. Longitudinal research is needed to determine whether XR-mediated increases in comprehension efficiency result in steady improvements in academic literacy, speaking, or writing. Comparative studies across different linguistic and educational contexts are also necessary to ascertain whether the benefits of XR are consistent across varying degrees of linguistic distance and institutional support, especially in formal EFL classrooms in Asia, the Middle East, or Latin America. Future research should therefore concentrate on adaptive XR designs that modify immersion intensity according to learner profiles (e.g., spatial ability, L1 background, and proficiency level). This will surpass universally applicable adaptive cognitive scaffolding. These enhancements would align MALL with embodied cognition principles and the practical realities of global language education, where equity and accessibility must remain top design priorities.

7 CONCLUSION

This study shows that it is possible to successfully develop consistent study habits among health sciences students, especially during the critical first year, by combining project-based learning with a mobile app-enhanced, data-informed approach. Students were inspired to study more often and were able to understand the value of consistency when daily app-based practice was incorporated into a cooperative group project. Individual effort was converted into group practice through the team discussions and shared goals, and the mobile app provided objective evidence of improvement. Students moved from intermittent, reactive studying to a more planned, contemplative, and sustained habit, which resulted in more than just more time spent studying alone. Academic achievement improved in measurable ways as a result of this model's consistency. It is a powerful illustration of the self-directed, life-long learning that is expected of health professionals after they graduate because of its simplicity and reality. The framework provides enough structure and social support to promote the early development of these habits. Additionally, it offers a practical and scalable approach to rethink how professional training may help students thrive right away.

8 DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

Microsoft Copilot was the only program we used to proofread the English text. The authors are responsible for the content and accuracy of the final version.

9 REFERENCES

[1] A. Al-Abri, F. Ranjbaran Madiseh, and M. Morady Moghaddam, "Exploring learning-oriented language assessment in enhancing students' lexical fluency through MALL," *The Asia-Pacific Educ. Res.*, vol. 34, pp. 1–13, 2024. https://doi.org/10.1007/s40299-024-00832-7

[2] K. Almudibry, "A meta-analysis of the literature on mobile-assisted language learning in response to COVID-19 in Saudi Arabia," *World J. Engl. Lang.*, vol. 12, no. 8, pp. 106–112, 2022. https://doi.org/10.5430/wjel.v12n8p106

[3] I. Barth, E. Spector-Cohen, R. Sitman, G. Jiang, F. Liu, and Y. Xu, "Beyond small chunks," *Int. J. Comput. Assist. Lang. Learn. Teach.*, vol. 9, pp. 79–97, 2019. https://doi.org/10.4018/IJCALLT.2019040105

[4] G. G. Botero, F. Questier, and C. Zhu, "Self-directed language learning in a mobile-assisted, out-of-class context: Do students walk the talk?" *Comput. Assist. Lang. Learn.*, vol. 32, pp. 71–97, 2018. https://doi.org/10.1080/09588221.2018.1485707

[5] F. Cakmak, "Mobile learning and mobile-assisted language learning in focus," *Lang. Technol.*, vol. 1, no. 1, pp. 30–48, 2019. https://dergipark.org.tr/en/pub/lantec/issue/42816/517381

[6] V. Chan, "Using a virtual reality mobile application for interpreting learning: Listening to the students' voice," *Interact. Learn. Environ.*, vol. 32, no. 6, pp. 2438–2451, 2024. https://doi.org/10.1080/10494820.2022.2147958

[7] N. P. Daly, "Investigating learner autonomy and vocabulary learning efficiency with MALL," *Lang. Learn. Technol.*, vol. 26, no. 1, pp. 1–30, 2022. https://doi.org/10.64152/10125/73469

[8] R. Dashtestani and S. Hojatpanah, "Mobile-assisted language learning in a secondary school in Iran: Discrepancy between the stakeholders' needs and the status quo," in *Handbook for Online Learning Contexts: Digital, Mobile and Open: Policy and Practice*, Springer, 2021, pp. 157–174. https://doi.org/10.1007/978-3-030-67349-9_12

[9] S. Ebadi and A. Raygan, "Investigating the facilitating conditions, perceived ease of use and usefulness of mobile-assisted language learning," *Smart Learn. Environ.*, vol. 10, no. 1, pp. 1–15, 2023. https://doi.org/10.1186/s40561-023-00250-0

[10] R. Fithriani, "The utilisation of mobile-assisted gamification for vocabulary learning: Its efficacy and perceived benefits," *CALL-EJ*, vol. 22, no. 3, pp. 146–163, 2021.

[11] I. García-Martínez, J. M. Fernández-Batanero, D. Cobos Sanchiz, and A. Luque de la Rosa, "Using mobile devices for improving learning outcomes and teachers' professionalisation," *Sustainability*, vol. 11, no. 24, p. 6917, 2019. https://doi.org/10.3390/su11246917

[12] N. Ghorbani and S. Ebadi, "Exploring learners' grammatical development in mobile-assisted language learning," *Cogent Educ.*, vol. 7, no. 1, p. 1704599, 2020. https://doi.org/10.1080/2331186X.2019.1704599

[13] T. Gonulal, "The use of Instagram as a mobile-assisted language learning tool," *Contemp. Educ. Technol.*, vol. 10, no. 3, pp. 309–323, 2019. https://doi.org/10.30935/cet.590108

[14] M. F. Hafour, "The effects of MALL training on preservice and in-service EFL teachers' perceptions and use of mobile technology," *ReCALL*, vol. 34, no. 3, pp. 274–290, 2022. https://doi.org/10.1017/S0958344022000015

[15] X. Lei, J. Fathi, S. Noorbakhsh, and M. Rahimi, "The impact of mobile-assisted language learning on English as a foreign language learners' vocabulary learning attitudes and self-regulatory capacity," *Front. Psychol.*, vol. 13, p. 872922, 2022. https://doi.org/10.3389/fpsyg.2022.872922

[16] R. Li, "Effects of mobile-assisted language learning on EFL/ESL reading comprehension," *Educ. Technol. Soc.*, vol. 25, no. 3, pp. 15–29, 2022. https://www.j-ets.net/ETS/journals/25_3/2.pdf

[17] D. Li, N. Tang, M. Chandler, and E. Nanni, "An optimal approach for predicting cognitive performance in education based on deep learning," *Comput. Hum. Behav.*, vol. 167, p. 108607, 2025. https://doi.org/10.1016/j.chb.2025.108607

[18] M. Menekse *et al.*, "Enhancing student reflections with natural language processing-based scaffolding: A quasi-experimental study in a large lecture course," *Comput. Educ.: Artif. Intell.*, vol. 8, p. 100397, 2025. https://doi.org/10.1016/j.caeai.2025.100397

[19] J. Park, J. Paxtle-Granjeno, M. W. Ok, M. Shin, and E. Wilson, "Preventing digital distraction in secondary classrooms: A quasi-experimental study," *Comput. Educ.*, vol. 227, p. 105223, 2025. https://doi.org/10.1016/j.compedu.2024.105223

[20] S. Pan, B. Hafez, A. Iskandar, and Z. Ming, "Integrating constructivist principles in an adaptive hybrid learning system for developing social entrepreneurship education among college students," *Learn. Motiv.*, vol. 87, p. 102023, 2024. https://doi.org/10.1016/j.lmot.2024.102023

[21] D. R. Isbell, H. Rawal, R. Oh, and S. Loewen, "Narrative perspectives on self-directed foreign language learning in a computer- and mobile-assisted language learning context," *Languages*, vol. 2, no. 2, p. 4, 2017. https://doi.org/10.3390/languages2020004

[22] L. Hsu, "Examining EFL teachers' technological pedagogical content knowledge and the adoption of mobile-assisted language learning: A partial least square approach," *Comput. Assist. Lang. Learn.*, vol. 29, no. 8, pp. 1287–1297, 2016. https://doi.org/10.1080/09588221.2016.1278024

[23] C. Troussas, A. Krouska, and M. Virvou, "Integrating an adjusted conversational agent into a mobile-assisted language learning application," in *Proc. IEEE 29th Int. Conf. Tools Artif. Intell. (ICTAI)*, 2017, pp. 1153–1157. https://doi.org/10.1109/ICTAI.2017.00179

[24] S. Tarighat and S. Khodabakhsh, "Mobile-assisted language assessment: Assessing speaking," *Comput. Hum. Behav.*, vol. 64, pp. 409–413, 2016. https://doi.org/10.1016/j.chb.2016.07.015

[25] I. Fronza and D. Gallo, "Towards mobile-assisted language learning based on computational thinking," in *Int. Conf. Web-Based Learn*, Cham: Springer, 2016, pp. 141–150. https://doi.org/10.1007/978-3-319-47440-3_16

[26] R. Shadiev, W. Y. Hwang, and Y. M. Huang, "Review of research on mobile language learning in authentic environments," *Comput. Assist. Lang. Learn*, vol. 30, nos. 3–4, pp. 284–303, 2017. https://doi.org/10.1080/09588221.2017.1308383

[27] H. Sudo, Y. Noborimoto, and J. Takahashi, "Grit, self-efficacy, and study habits in mobile learning among physical therapy students," *International Journal for Educational Media and Technology*, vol. 19, no. 1, 2025.

[28] B. Moreillon *et al.*, "Prediction of plasma volume and total hemoglobin mass with machine learning," *Physiological Reports*, vol. 11, no. 19, p. e15834, 2023. https://doi.org/10.14814/phy2.15834

10 AUTHORS

Antony Desilva D. is a research scholar at the Department of English at B.S. Abdur Rahman Crescent Institute of Science and Technology, India. His research area is Neurolinguistic Programming.

Dr. Vijayakumar Selvaraj is an Associate Professor in the Department of English at B.S. Abdur Rahman Crescent Institute of Science and Technology, India, with a PhD in Applied Linguistics. His research is anchored in computer-assisted language learning, AI-driven pedagogical innovation, and inclusive digital education. Over 18 publications have appeared in Q1 journals such as The International Review of Research in Open and Distributed Learning, Humanities & Social Sciences Communications, and Transactions on Emerging Telecommunications Technologies. His interdisciplinary work integrates natural language processing, neuro-linguistic programming, and adaptive learning systems to enhance L2 proficiency, accessibility, and equity in digital learning environments. Several patents have been filed on AI-based educational tools, including assistive technologies for visually impaired learners. He supervises six doctoral candidates and serves as a peer reviewer for multiple Q1 journals. His scholarly output reflects a sustained commitment to evidence-based interventions that bridge pedagogy, technology, and linguistic equity (E-mail: vijayakumar@crescent.education).

Dr. Sathikulameen A. is an Assistant Professor and Research Supervisor in the Postgraduate & Research Department of English at The New College, Chennai. With 16 years of experience in education, he specializes in English Language Teaching, Language Acquisition and Pedagogy, Multimedia and Technology-Enhanced Language Learning, and Cultural Studies. He has published 18 research papers in various indexed journals, contributed to conference proceedings, and authored several books and chapters in edited volumes. He has also been actively involved in organizing and coordinating academic programs, workshops, and conferences. He has served as a resource person, keynote speaker, and examiner in numerous academic and professional development events.

Emmanuel Rajkumar B. is a research scholar at the Postgraduate and Research Department of English at The New College, University of Madras. His research areas include Ecocriticism, Cultural Studies, and English Language Teaching.

Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 4 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i04.60519

PAPER

Immersive Music Therapy Using Virtual Reality and Mobile Technologies

Sai Wang (✉)

Hebei University of Environmental Engineering, Qinhuangdao, China

18233545656@163.com

ABSTRACT

The global prevalence of mental disorders such as depression and anxiety continues to rise. Traditional music therapy faces challenges such as a lack of immersion, delayed emotional response, and rigid rehabilitation pathways, making it difficult to meet the demand for personalized treatment. The immersive characteristics of virtual reality (VR) and the portability of interactive mobile technologies provide technical support to overcome these limitations, with accurate emotion recognition being a key prerequisite for personalized intervention. This paper aims to construct an immersive music therapy environment based on the deep integration of VR and interactive mobile technologies and proposes a dual-level fusion method for multimodal emotion recognition, combining feature-level and model-level integration to dynamically optimize the psychological rehabilitation path. Methodologically, we first design a "hardware collaboration-software adaptation-interactive feedback loop" architecture that integrates physiological signal acquisition, VR scene rendering, and mobile interaction control modules. Multimodal data, including EEG, ECG, facial expression images, and subjective emotional ratings, are collected. These are then aligned and weighted across modalities at the feature level to extract high-level features, and deep learning models with attention mechanisms at the model level are used for precise emotion classification. Finally, based on real-time emotion recognition results, an optimization algorithm driven by reinforcement learning is developed to dynamically adjust music parameters and VR scene elements. The study confirms that the integration of VR and interactive mobile technologies can break through the limitations of traditional therapy scenarios. The dual-level fusion strategy provides higher accuracy and robustness for emotion recognition, while the dynamic optimization mechanism offers personalized solutions for psychological rehabilitation, with significant academic innovation and clinical application potential.

KEYWORDS

virtual reality (VR), interactive mobile technologies, immersive music therapy, multimodal emotion recognition, feature-level fusion, model-level fusion, psychological rehabilitation path optimization

Wang, S. (2026). Immersive Music Therapy Using Virtual Reality and Mobile Technologies. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(4), pp. 120–134. https://doi.org/10.3991/ijim.v20i04.60519

Article submitted 2025-10-22. Revision uploaded 2025-12-12. Final acceptance 2022-12-14.

1 INTRODUCTION

The global mental health crisis continues to worsen. According to a report from the World Health Organization (WHO), approximately 280 million people worldwide suffer from depression, and 260 million people are affected by anxiety disorders [1, 2]. These mental health disorders have become one of the leading causes of disability-adjusted life years lost globally, resulting in social and economic losses amounting to trillions of dollars each year [3, 4]. Against the backdrop of increasing life pressures, the imbalance between the supply and demand for traditional psychological treatment resources has become more apparent. Developing efficient, convenient, and personalized treatment technologies has become an urgent need in the global public health field [5, 6]. Music therapy, as a core non-pharmacological intervention, is widely applied in psychological rehabilitation. However, traditional models face significant bottlenecks: treatment scenarios are limited to fixed clinics, making it difficult to create a deeply immersive emotional resonance environment [7]; emotional assessments rely on patient self-reports and therapists' experiences, which are delayed and highly subjective [8]; rehabilitation pathways are based on general guidelines, lacking individual dynamic adaptability, which severely restricts their clinical effectiveness and promotion [9].

The immersive characteristics of virtual reality (VR) can create highly realistic treatment scenarios and reduce patients' psychological defenses [10]; interactive mobile technologies break spatial limitations, enabling portability and real-time treatment [11]; multimodal emotion recognition integrates objective data to provide core support for personalized treatment. The fusion of these three elements offers a new path to overcome these bottlenecks [12]. However, existing research has clear deficiencies: the integration of VR and mobile technologies is mostly functional addition, lacking the "immersion-interaction-data" closed-loop architecture [13]; feature-level fusion in multimodal emotion recognition faces the challenge of data heterogeneity, while model-level fusion is highly complex and lacks adaptation to patients with psychological disorders [14]; rehabilitation path optimization relies on historical data, lacking real-time emotional feedback, and core intervention variable regulation is insufficiently refined [15]. Currently, the field has yet to form an integrated solution of "VR-interactive mobile fusion environment + high-precision multimodal emotion recognition + dynamic rehabilitation path optimization," which provides the core entry point for this study.

This study sets three main objectives: ① to construct a VR-interactive mobile fusion music therapy environment with both high immersion and portability; ② to propose a feature-level and model-level dual-fusion multimodal emotion recognition method; ③ to design a rehabilitation path dynamic optimization mechanism based on real-time emotional feedback. The innovations are reflected in three aspects: ① Architectural innovation, achieving the organic unity of immersion experience and mobile control through device protocol adaptation and data synchronization; ② Methodological innovation, using cross-modal feature alignment and attention-weighted fusion strategies to improve emotion recognition accuracy and robustness; ③ Mechanism innovation, constructing a closed-loop optimization model driven by reinforcement learning to dynamically adjust treatment parameters and achieve personalized rehabilitation.

The structure of the paper is as follows: Chapter 2 describes the architecture design and core module implementation of the immersive treatment environment; Chapter 3 details the dual-level fusion method for multimodal emotion recognition; Chapter 4 builds a reinforcement learning-driven rehabilitation path optimization mechanism; Chapter 5 verifies the effectiveness of the solution through clinical experiments; Chapter 6 discusses the research value, limitations, and future directions; Chapter 7 summarizes the core findings and prospects for application.

2 CONSTRUCTION OF IMMERSIVE MUSIC THERAPY ENVIRONMENT

This study is based on the hierarchical architecture model of multimodal interaction and rehabilitation path optimization in the immersive music therapy environment shown in Figure 1. A six-level collaborative architecture is designed: Technology Layer – Perception Interaction Layer – Data Processing Layer – Algorithm Decision Layer – Application Layer – Feedback Optimization Layer. This architecture achieves the organic integration of immersive experience and real-time interaction through precise hardware and software adaptation. The technology layer adopts a 5G + Wi-Fi 6 dual-mode transmission protocol, and the hardware integrates the Oculus Quest 2 VR headset, Android/iOS mobile terminals, and multimodal physiological data collection devices. The software relies on Unity3D, Flutter, and microservice cloud platforms for cross-terminal collaboration. The perception interaction layer synchronously collects physiological signals such as electroencephalography (EEG), electrocardiography (ECG), skin conductance, and facial expression behavior characteristics. It also enables bi-directional interaction of scene interaction and parameter adjustment through VR headsets, motion controllers, and mobile terminals. The data processing layer is deployed on edge cloud nodes and provides high-quality data for subsequent analysis through filtering, ICA denoising, and normalization algorithms. The algorithm decision layer adopts a feature-level and model-level dual-fusion multimodal emotion recognition method, combined with reinforcement learning algorithms, to dynamically output optimized music parameters, VR scenes, and interaction tasks. The application layer provides treatment plan recommendations, multimodal data visualization, and HL7 standard data interface functions. The feedback optimization layer closes the feedback loop of algorithm decision results to the interaction and application layers while continuously optimizing the intervention strategy based on the patient's subjective Self-Assessment Manikin (SAM) scores. This six-level architecture forms a complete closed loop of Collection – Transmission – Processing – Decision – Application – Feedback, providing systemic support for the technical implementation and clinical application of immersive music therapy.

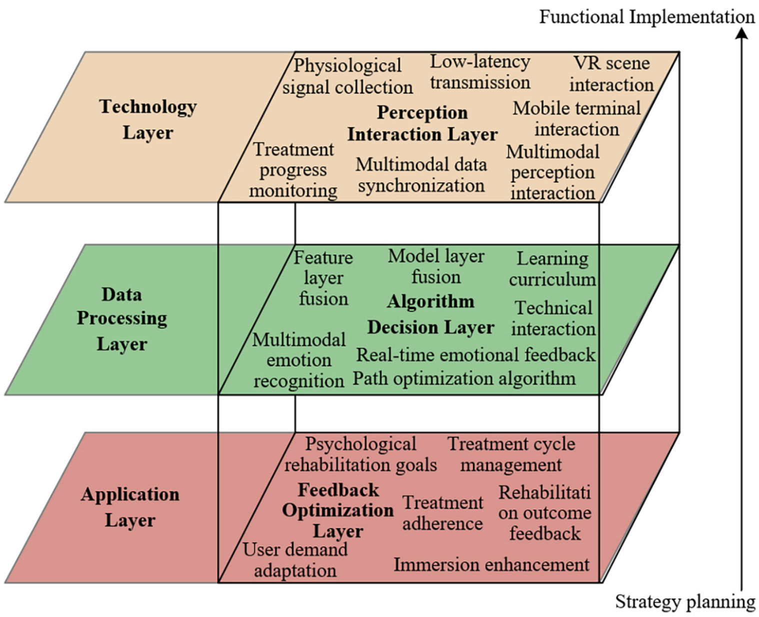


Fig. 1. Hierarchical architecture model of multimodal interaction and rehabilitation path optimization in immersive music therapy environment

Core modules collaborate to ensure the immersion and reliability of treatment. VR scenes are designed with clinical requirements for both relaxing and stimulating systems. Environmental psychology is referenced to optimize scene parameters, and detail-level technologies are used to improve rendering efficiency. A field-of-view adaptive adjustment is applied to control the incidence of dizziness below 5%. The music library integrates international therapeutic music resources and establishes a database of parameters such as rhythm, pitch, and volume. Scene-music mapping algorithms are used to ensure smooth transitions when switching between parameters. The core functions of the mobile app include intelligent treatment plan recommendations, real-time display of emotional and physiological indicators, and slider-based parameter adjustment, compatible with HL7 standards for data integration with clinical systems. Physiological signal preprocessing targets noise removal: a 50Hz notch filter is used to eliminate power line interference, a 4–30 Hz low-pass filter extracts effective EEG signals, independent component analysis (ICA) separates eye movement and muscle artifacts, and the data is finally normalized using Z-score normalization, mapping multimodal data to the same feature space to provide high-quality input for subsequent emotion recognition.

3 MULTIMODAL EMOTION RECOGNITION METHOD

3.1 Emotion modal data collection plan

The emotion modal data collection in this study strictly follows clinical experimental protocols and data standardization principles. The experimental subjects are patients diagnosed with depression/anxiety disorder according to the diagnostic criteria of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-5*). Inclusion criteria: age 18–55 years, primary school education or above, and ability to cooperate with the collection process. Exclusion criteria: comorbidity of severe physical diseases, history of epilepsy, allergy to VR devices, or cognitive dysfunction. The data collection process is divided into three stages: the baseline testing stage, where the participant's basic physiological signals and facial expression data are collected for five minutes in a quiet environment, along with completing a basic information questionnaire; the emotion induction stage, where different valence images from the International Affective Picture System are selected and combined with a self-constructed standardized emotional music library, presented through an immersive VR scene to induce the target emotion; and the data recording stage, where multimodal data is continuously collected until the presentation of the inducing materials is complete. The collected data includes three types: physiological data collected using specialized equipment, EEG recorded through a 16-channel amplifier at key electrode points in the international 10–20 system (Fp1, Fp2, etc.), ECG to record heart rate, RR interval, and other indicators, and skin conductance response to record peak and rise slope; behavioral data captured by the VR headset's built-in high-definition camera to capture facial expression images, with 68 facial feature points extracted based on the Dlib library; subjective data collected through a self-assessment model, where the participant rates each emotional induction material on a 1–9 scale for pleasure and arousal levels, providing complementary verification of objective data and subjective experiences.

3.2 Feature-level fusion strategy

The goal of single-modal feature extraction is to accurately capture the emotional representations of each modality. Differentiated extraction schemes are designed for different data types. In terms of physiological features, EEG signals extract time-domain and frequency-domain features based on preprocessed data. Time-domain features, such as mean, standard deviation, and kurtosis, reflect the signal amplitude distribution characteristics. Frequency-domain features, calculated through fast Fourier transform, include power spectral density of α waves (8–13Hz), β waves (14–30Hz), and θ waves (4–7Hz), representing the brain's electrical activity rhythms under different emotional states. ECG features focus on heart rate and heart rate variability indicators, extracting core parameters such as the mean RR interval, standard deviation, and root mean square of adjacent RR interval differences. GSR features extract signal peak values, rising slopes, and mean amplitudes, reflecting skin conductance responses triggered by emotional arousal levels. In terms of behavioral features, a lightweight CNN architecture is used to extract features from preprocessed facial expression images. Depth wise separable convolutions are employed to reduce the model parameters while retaining key facial expression features, ultimately outputting a 256-dimensional facial depth feature vector. Subjective features directly quantify the pleasure and arousal ratings from the self-assessment model into a 2-dimensional feature vector, achieving numerical representation of subjective emotional experience.

The focus of feature-level fusion is to address the data heterogeneity problem of multimodal data and construct high-level feature representations for cross-modal collaboration. The first step is to perform feature alignment: for time-series physiological data like EEG and ECG, frame-level synchronization is performed based on timestamps; all single-modal features are Z-score normalized, mapping feature values to the [0,1] range to eliminate dimensional differences. Based on this, a dynamic weighted fusion strategy is designed. The contribution of each single-modal feature to emotion recognition is calculated through a validation set experiment, and higher weights are assigned to modalities with higher contributions. The specific fusion formula is defined as follows:

$$V = \omega_1 E_{EEG} + \omega_2 E_{ECG} + \omega_3 E_{GSR} + \omega_4 E_{Face} + \omega_5 E_{SAM} \tag{1}$$

where, V is the fused feature vector, ω_i is the weight of the i-th single-modal feature, and F_{EEG}, F_{ECG}, ... correspond to each single-modal feature vector. This strategy not only avoids information redundancy caused by simple feature concatenation but also strengthens the representation ability of effective information through dynamic weight allocation, providing high-recognition multimodal base features for subsequent model-level fusion.

3.3 Model-level fusion strategy

To address the limitations of feature-level fusion in mining complex intermodal associations and eliminating redundant information, this study constructs a multi-model ensemble architecture based on a cross-modal attention mechanism, achieving deep collaborative learning and precise representation of multimodal emotional information. Based on the characteristics of different modal data, a heterogeneous base model system is designed: physiological features with strong temporal dependence are input to a bidirectional long short-term memory network (BiLSTM) to capture the temporal evolution patterns of emotional states; deep facial expression features are input to a lightweight convolutional neural network (CNN) model to further extract local key emotional representations; and subjective features, along with intermediate features from the above modalities, are integrated and input to a multi-layer perceptron (MLP) to adapt to the nonlinear mapping requirements of high-dimensional heterogeneous data. The core innovation lies in the introduction of a cross-modal attention mechanism. By constructing an attention weight matrix, the importance coefficient of each base model's output result is dynamically learned. For example, in anxiety emotion recognition, higher weight is assigned to the EEG feature model, which contributes more, while also uncovering the potential intermodal associations, achieving adaptive fusion of multi-model outputs and generating a more compact and information-rich shared emotional feature vector.

To ensure the generalization ability and classification performance of the model, a multi-dimensional regularization collaborative optimization mechanism is established to effectively suppress overfitting. During model training, L_2 regularization is applied to penalize the weight parameters of each base model, reducing model complexity by adding a weight squared term to the loss function. The weight squared term expression is given by:

$$L_{total} = L_{CE} + \lambda \sum \|W\|_2^2 \tag{2}$$

where L_{CE} is the cross-entropy loss, λ is the regularization coefficient, and W is the model weight matrix. Dropout layers are embedded in both CNN and MLP layers, with a deactivation probability of 0.2 to randomly mask some neurons and avoid over-reliance on local features. Additionally, early stopping is employed to monitor the validation set loss, and training is terminated when the loss does not decrease for 10 consecutive training epochs, retaining the model parameters with the best generalization performance. The optimizer used for model optimization is Adaptive Moment Estimation (Adam), with an initial learning rate of 0.001, and a cosine annealing strategy is used to dynamically adjust the learning rate, improving training stability and convergence speed. Finally, the fused shared features are input into a Softmax classifier, which outputs the probability distribution of four emotional states: joy, calmness, anxiety, and depression, completing the emotion recognition task.

$$V = \sum_i \omega_i F_i$$

$$Loss = L_{CE} + \lambda \|W\|^2$$

4 PSYCHOLOGICAL REHABILITATION PATH OPTIMIZATION

The precise construction of the rehabilitation path is based on the dissection of core elements and scientific initial configuration, laying the foundation for subsequent dynamic optimization. This study identifies five core elements of the rehabilitation path, forming a multidimensional collaborative intervention system: the treatment goal distinguishes between short-term and long-term levels, with the short-term focusing on immediate relief of anxiety/depression emotions and the long-term aiming for comprehensive restoration of psychosocial functions; the treatment cycle is set to 8–12 weeks according to clinical guidelines, with two treatments per week, each lasting 45 minutes; the music parameter combinations cover three core dimensions: rhythm, pitch, and volume; VR scene types are adapted to the treatment stages, initially using a relaxing natural scene and switching to an encouraging dynamic scene based on emotional improvement; interactive tasks include lightweight tasks such as scene exploration and virtual instrument playing to enhance patient engagement. The initial path generation relies on the core recommendations from the *China Depression Disorder Prevention and Treatment Guidelines* (2023 Edition) and the *Anxiety Disorder Diagnosis and Treatment Guidelines* (2021 Edition), combined with the consensus experience of three senior experts in psychotherapy. The initial path is divided into three levels—mild, moderate, and severe—based on the patient's initial Self-Rating Depression Scale (SDS)/Self-Rating Anxiety Scale (SAS) scores. A differentiated initial path library is constructed; for example, in the case of severe patients, the initial path focuses on low-stimulation relaxing music and forest VR scenes, gradually increasing the intervention intensity.

The reinforcement learning-driven optimization model provides intelligent technical support for the dynamic adjustment of the rehabilitation path. Its core lies in achieving adaptive iteration of the path through environmental interaction and reward feedback. This study deeply binds the core components of reinforcement learning with the treatment scene: the agent is defined as the immersive music therapy system, responsible for perceiving the patient's state and executing path adjustment actions; the environment includes the VR immersive scene and the patient's real-time state, forming a dynamic interaction carrier; the state space is represented by a high-dimensional vector, including key physiological indicators such as multimodal emotion recognition results, treatment progress ratio, real-time SDS/SAS scores, and heart rate variability, forming a 32-dimensional state vector; the action space focuses on fine-tuning the three main intervention variables, with music parameters supporting rhythm adjustments of ±5 BPM, pitch adjustments of 1 scale, and volume adjustments of ±5 dB, while VR scenes provide four scene switching options, and interactive tasks include updates to three difficulty levels; the reward function design adopts a deviation quantification mechanism, defined as:

$$R = 1 - \left| S_{real} - S_{tar} \right| / S_{\max} \tag{3}$$

where S_{real} is the real-time emotional scale score of the patient, S_{tar} is the target score for the corresponding treatment stage, and S_{max} is the maximum score. When S_{real} approaches S_{tar}, the reward value approaches 1; otherwise, it decreases. The optimization algorithm uses Deep Q-Network (DQN) to build a "convolutional layer-fully connected layer" neural network architecture, with the convolutional layer extracting high-dimensional features from the state space. The fully connected layer fits the action value function, and the experience replay mechanism stores historical interaction data and randomly samples for training. The target network periodically updates the policy, ensuring the model converges to the optimal treatment path.

The dynamic execution mechanism ensures the effectiveness and individual adaptation of the optimized path through real-time feedback loops and personalized adjustments. This study designs a four-stage closed-loop execution process of "Emotion Recognition – Path Evaluation – Parameter Adjustment – Effect Monitoring," setting every two treatment cycles as an adjustment window. First, based on the multimodal emotion recognition results and SDS/SAS retest data, the effectiveness of the current path is evaluated. Then, the reinforcement learning model generates the optimal adjustment plan, simultaneously updating the music parameters, VR scenes, and interactive tasks. The adjusted path is executed in subsequent treatments, while real-time monitoring of the patient's physiological indicators and emotional state provides data support for the next round of adjustments. To account for individual patient differences, an individual difference factor γ is introduced to dynamically optimize the weight distribution of the reward function. For elderly patients or those with a long course of illness, γ is set to a lower value to slow down the adjustment rate of intervention intensity, while for younger or more responsive patients, the γ value is increased to accelerate the path optimization. This mechanism achieves the precise transformation of the rehabilitation path from "generalized" to "personalized," ensuring that the intervention strategy highly matches the individual characteristics of the patient and the treatment dynamics. Figure 2 presents the multimodal data interaction and rehabilitation path optimization process flowchart for the immersive music therapy environment, visually demonstrating the complete closed-loop process of "Patient Treatment Interaction – Multimodal Data Collection – Emotion Recognition and Path Optimization – Treatment Plan Adjustment Feedback."

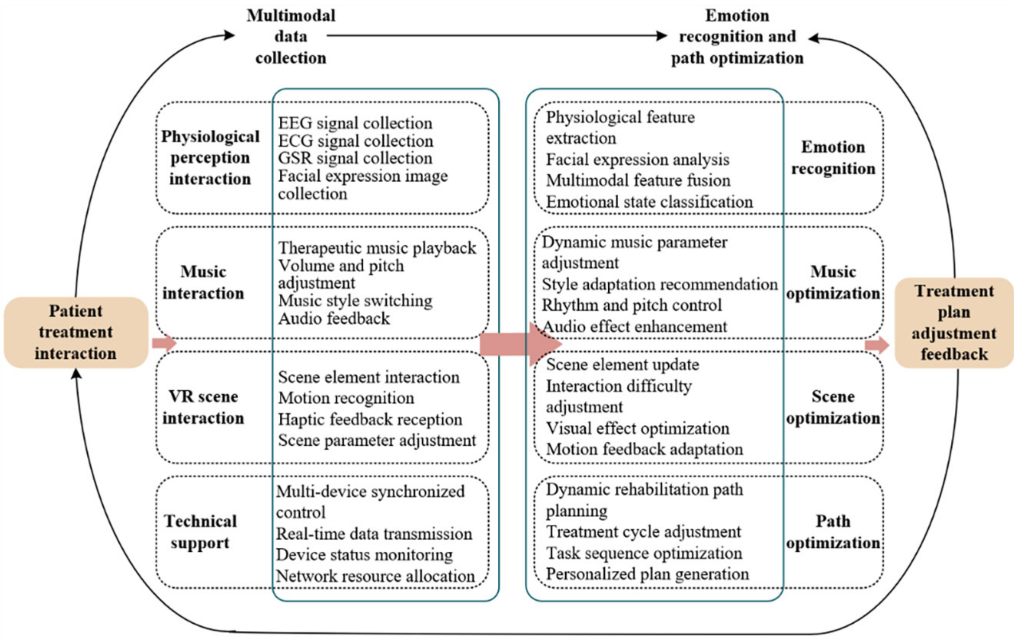


Fig. 2. Multimodal data interaction and rehabilitation path optimization process flowchart for immersive music therapy environment

5 EXPERIMENT DESIGN AND RESULT ANALYSIS

5.1 Experiment design

This experiment strictly follows the protocol of a randomized controlled clinical trial. A total of 120 patients diagnosed with depression/anxiety were selected as experimental subjects, with the sample size determined using GPower3.1 software. The inclusion criteria are as follows: meeting the diagnostic criteria of the *DSM-5*, SDS score ≥ 53 or SAS score ≥ 50, age between 18–60 years, ability to operate

VR devices, and voluntary signing of the informed consent form. Exclusion criteria include the presence of severe physical illnesses, a history of epilepsy, cognitive dysfunction, or having received other psychological interventions in the past three months. Patients were randomly assigned to the experimental group and the control group, each with 60 individuals, using a random number table. Baseline data comparison showed no significant differences between the two groups in terms of age, gender ratio, disease duration, and initial SDS/SAS scores ($P > 0.05$), indicating comparability between the groups. In terms of equipment, the experimental group used the Oculus Quest 3 VR headset, a 16-channel EEG amplifier, a portable ECG sensor, and a self-developed therapeutic interactive app, with data processing supported by Alibaba Cloud's edge computing nodes. The control group used a professional audio system and a tablet with a standard music library installed. The experimental treatment cycle lasted for 10 weeks, with both groups undergoing treatment three times a week, each session lasting 45 minutes. Data collection occurred before treatment, every two weeks during treatment, and after treatment in stages. Equipment calibration was completed before treatment, real-time operational status was recorded during treatment, and subjective rating data was collected post-treatment.

5.2 Experiment results analysis

Table 1. Immersive treatment environment performance test results

Evaluation Dimension	Specific Indicator	Unit	Experimental Group Test Value	Industry Reference Standard	Compliance Status
Immersion	IPQ Scale – Spatial Presence	Points (1–5)	4.2 ± 0.3	≥3.5 Points	Compliant
	IPQ Scale – Sense of Reality Loss	Points (1–5)	4.0 ± 0.4	≥3.0 Points	Compliant
	IPQ Scale – Involvement	Points (1–5)	4.3 ± 0.3	≥3.5 Points	Compliant
	IPQ Total Score	Points (1–5)	4.2 ± 0.3	≥3.5 Points	Compliant
Real-Time Performance	EEG Signal Collection Delay	*ms*	12.3 ± 1.5	≤20 *ms*	Compliant
	Multimodal Data Transmission Delay	*ms*	28.5 ± 2.1	≤50 *ms*	Compliant
	VR Scene Switch Delay	*ms*	35.2 ± 3.0	≤50 *ms*	Compliant
	Music Parameter Adjustment Delay	*ms*	18.7 ± 1.8	≤30 *ms*	Compliant
Stability	Continuous Operation Time	*h*	72	≥48 *h*	Compliant
	Device Failure Rate	%	0.8	≤2%	Compliant
	Data Loss Rate	%	0.3	≤1%	Compliant
User Adaptability	Dizziness Rate	%	4.2	≤10%	Compliant

To verify whether the constructed VR-interactive mobile integrated treatment environment meets clinical needs in terms of immersion, real-time performance, and stability, a performance test experiment was conducted. As shown in Table 1, the experimental group exceeded industry reference standards in all core performance indicators:

- Immersion: The IPQ scale scores in the three dimensions of spatial presence, sense of reality loss, and involvement, as well as the total score, all exceeded the qualified threshold of 3.5 points. Specifically, spatial presence and involvement scored above 4.0 points, indicating that the VR scene effectively creates an immersive therapeutic atmosphere.
- Real-Time Performance: Key processes, including EEG signal collection and data transmission delays, were controlled within 50ms. Music parameter adjustments and scene switching had delays of 18.7ms and 35.2ms, respectively, meeting the technical requirements for real-time emotional feedback and dynamic path adjustments.
- Stability: The device was continuously operated for 72 hours, with a failure rate of only 0.8% and a data loss rate of 0.3%. The dizziness rate was as low as 4.2%, demonstrating the clinical applicability of the environment.

In conclusion, the core performance of the constructed environment meets the standards, providing reliable technical support for the subsequent implementation of personalized treatments.

Table 2. Performance comparison of different emotion recognition methods

Recognition Method		Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	AUC Value
Single Modality Methods	EEG Recognition Alone	72.3 ± 1.5	71.8 ± 1.7	70.5 ± 1.6	71.1 ± 1.5	0.782
	Facial Expression Recognition Alone	75.6 ± 1.3	74.9 ± 1.4	73.2 ± 1.5	74.0 ± 1.4	0.805
Single Fusion Methods	Feature Layer Fusion Only	83.5 ± 1.2	82.9 ± 1.3	81.7 ± 1.2	82.3 ± 1.2	0.868
	Model Layer Fusion Only	84.2 ± 1.1	83.5 ± 1.2	82.4 ± 1.1	82.9 ± 1.1	0.875
	91.6 ± 0.8	91.6 ± 0.8	90.8 ± 0.9	90.2 ± 0.8	90.5 ± 0.8	0.943
Ablation Experiments	Removing Feature Layer Fusion	80.3 ± 1.4	79.6 ± 1.5	78.5 ± 1.4	79.0 ± 1.4	0.841
	Removing Model Layer Fusion	79.5 ± 1.3	78.9 ± 1.4	77.8 ± 1.3	78.3 ± 1.3	0.836

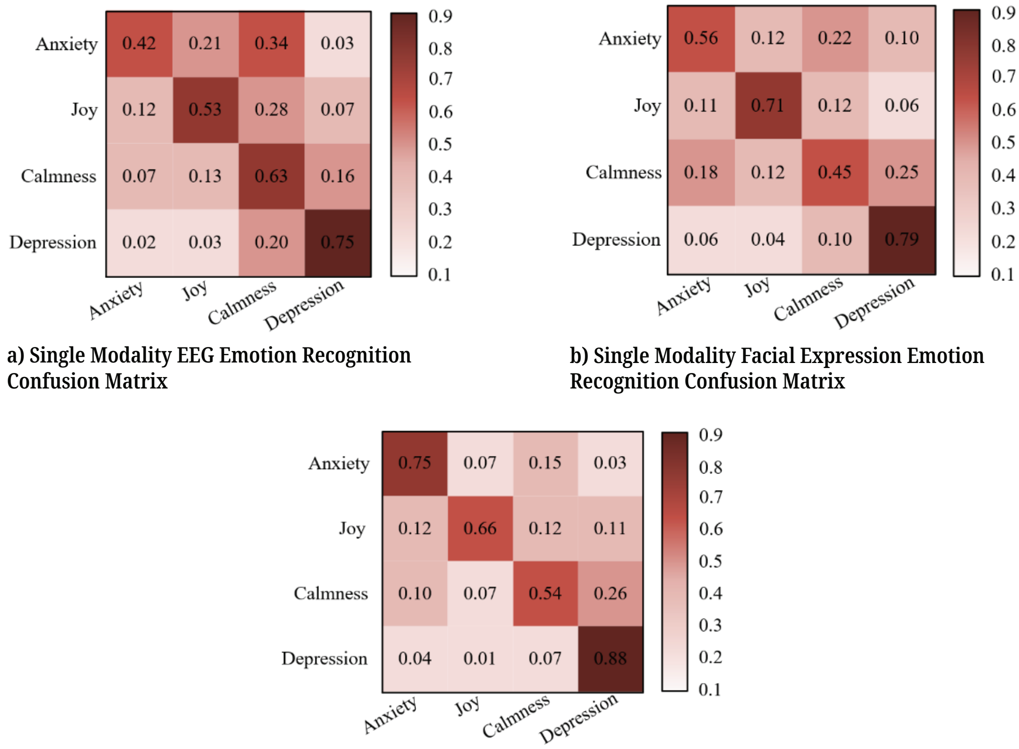
To verify the recognition performance of the proposed feature-layer-model-layer two-level fusion method and the necessity of the two-level fusion strategy, comparative and ablation experiments were conducted. As shown in Table 2, the proposed method significantly outperformed the single-modality and single-fusion methods in all core indicators: the accuracy reached 91.6%, which is an improvement of 16.0 percentage points compared to the best single-modality method and

7.4 percentage points compared to the best single-fusion method (model layer fusion); the F1 score and AUC value reached 90.5% and 0.943, respectively, reflecting the method's advantage in classification accuracy and generalization ability. The ablation experiment results showed that removing either the feature-layer or model-layer fusion resulted in a noticeable decrease in recognition performance. The accuracy dropped to 80.3% and 79.5%, and the F1 score dropped by more than 11 percentage points, indicating that the cross-modality information integration at the feature layer and the attention-weighted integration at the model layer form a synergistic effect, both of which are essential. In conclusion, the proposed two-level fusion method can achieve accurate emotion state recognition for patients with psychological disorders and provide reliable data input for dynamic optimization of rehabilitation pathways.

Table 3. Comparison of rehabilitation effect scores between the experimental group and the control group (x ± s, points)

Evaluation Indicator	Group	Pretreatment	4 Weeks of Treatment	8 Weeks of Treatment	10 Weeks of Treatment	Change Rate from Baseline After Treatment	Intergroup Difference (P Value)
Depression Scale (SDS)	Experimental Group	62.5 ± 7.3	53.2 ± 6.8	45.1 ± 5.9	40.2 ± 5.2	–35.7%	<0.01
	Control Group	63.1 ± 7.5	58.6 ± 7.1	52.3 ± 6.5	48.5 ± 6.1	–23.1%	
Anxiety Scale (SAS)	Experimental Group	58.6 ± 6.9	49.5 ± 6.2	42.3 ± 5.5	37.8 ± 4.9	–35.5%	<0.01
	Control Group	59.2 ± 7.2	54.8 ± 6.8	49.1 ± 6.2	45.3 ± 5.8	–23.5%	
Quality of Life Scale – Physical Dimension	Experimental Group	58.2 ± 6.5	64.5 ± 6.1	70.3 ± 5.8	75.6 ± 5.3	+29.9%	<0.01
	Control Group	57.8 ± 6.7	60.2 ± 6.3	63.5 ± 5.9	66.8 ± 5.6	+15.6%	
Quality of Life Scale – Psychological Dimension	Experimental Group	55.3 ± 7.1	62.1 ± 6.5	68.4 ± 6.0	74.2 ± 5.5	+34.2%	<0.01
	Control Group	54.9 ± 7.3	58.6 ± 6.8	62.3 ± 6.2	65.1 ± 5.8	+18.6%	
Quality of Life Scale – Social Relationship Dimension	Experimental Group	59.1 ± 6.8	65.3 ± 6.2	71.5 ± 5.9	76.8 ± 5.4	+29.9%	<0.01
	Control Group	58.7 ± 7.0	61.5 ± 6.5	65.2 ± 6.0	68.9 ± 5.7	+17.4%	
Quality of Life Scale – Environmental Dimension	Experimental Group	60.5 ± 6.6	66.8 ± 6.1	72.4 ± 5.8	77.3 ± 5.2	+27.8%	<0.01
	Control Group	60.1 ± 6.8	63.2 ± 6.3	66.5 ± 5.9	69.8 ± 5.5	+16.1%	

To verify the clinical therapeutic advantages of the optimized rehabilitation path, a comparison of the rehabilitation effect differences between the experimental group and the control group was conducted. As shown in Table 3, both groups showed a reduction in SDS and SAS scores and an increase in quality-of-life scale scores after treatment. However, the experimental group showed significantly better improvements than the control group: after 10 weeks of treatment, the experimental group's SDS and SAS scores decreased by 35.7% and 35.5%, respectively, which was more than 12 percentage points higher than the control group; the total quality of life scale score and improvements in the physical, psychological, social relationship, and environmental dimensions all exceeded 27%, which was 10–16 percentage points higher than the control group. Statistical analysis showed that the experimental group showed significant improvements after four weeks of treatment ($P < 0.05$), and the trend of improvement continued to strengthen. After 10 weeks of treatment, the intergroup differences reached a significant level ($P < 0.01$). This indicates that the dynamically optimized rehabilitation path based on real-time emotional feedback can more accurately adapt to patients' emotional changes. Combined with the immersive advantages of the VR-mobile integrated environment, it significantly enhances the improvement of depression, anxiety symptoms, and quality of life, with notable clinical therapeutic value.



a) Single Modality EEG Emotion Recognition Confusion Matrix

b) Single Modality Facial Expression Emotion Recognition Confusion Matrix

c) Proposed Two-Level Fusion Emotion Recognition Confusion Matrix

Fig. 3. Performance comparison of multimodal emotion recognition methods' confusion matrices

To verify the performance advantages of the proposed feature-layer-model-layer two-level fusion method in emotion recognition for patients with psychological disorders, this study compared the confusion matrices of single-modality and two-level fusion methods. From the results in Figure 3, it can be seen that in single-modality EEG recognition, the accuracy of depression recognition reached 0.75, but anxiety was only 0.42, indicating a clear category bias. In single-modality facial expression recognition, the recognition accuracy for joy was 0.71, and depression was 0.79, but the cross-category discriminability for anxiety (0.56) and calmness (0.45) was insufficient. In the proposed two-level fusion method, the recognition accuracy for anxiety (0.75), joy (0.66), calmness (0.54), and depression (0.88) all significantly improved, and the balance of recognition accuracy for each emotion category and the overall accuracy far exceeded the single-modality methods. This indicates that the two-level fusion strategy effectively integrates multimodal physiological and behavioral information, overcoming the dimensional limitations of single-modality data, and enables accurate recognition of complex emotional states in psychological disorder patients. This provides high-confidence data input for dynamic optimization of rehabilitation paths, fully demonstrating the method's technical adaptability and clinical value in psychological rehabilitation scenarios.

Table 4. User experience evaluation results of the experimental group

Evaluation Dimension	Specific Indicator	Unit	Test Value	Reference Standard	Compliance Status
Immersion Experience	IPQ Scale – Spatial Presence	Points (1–5)	4.2 ± 0.3	≥ 3.5 points	Compliant
	IPQ Scale – Loss of Reality	Points (1–5)	4.0 ± 0.4	≥ 3.0 points	Compliant
	IPQ Scale – Involvement	Points (1–5)	4.3 ± 0.3	≥ 3.5 points	Compliant
	IPQ Scale – Total Score	Points (1–5)	4.2 ± 0.3	≥ 3.5 points	Compliant
Treatment Adherence	Treatment Attendance Rate	%	96.2	≥80%	Excellent
	Interaction Task Completion Rate	%	94.5	≥85%	Excellent
	Treatment Plan Execution Rate	%	93.8	≥85%	Excellent
Device and Function Satisfaction	Device Operation Convenience Rating	Points (1–10)	8.6 ± 1.2	≥ 7 points	Satisfied
	Function Adaptability Rating	Points (1–10)	8.4 ± 1.3	≥ 7 points	Satisfied
	Overall Satisfaction Rating	Points (1–10)	8.5 ± 1.2	≥ 7 points	Satisfied
Adverse Reactions	Device Discomfort Incidence Rate	%	5.3	≤10%	Compliant
Immersion Experience	IPQ Scale – Spatial Presence	Unit	Test Value	Reference Standard	Compliance Status

To verify the user acceptance and clinical feasibility of the immersive treatment environment, a comprehensive evaluation of the user experience in the experimental group was conducted. As shown in Table 4, the experimental group performed excellently in core dimensions such as immersion, treatment adherence, and device satisfaction: The scores for the three dimensions of the IPQ scale and the total score all exceeded the passing threshold of 3.5, with spatial presence and involvement reaching scores above 4.0, indicating that the VR scene's immersive experience effectively enhances emotional involvement. The treatment attendance rate, interaction task completion rate, and plan execution rate all exceeded 93%, far surpassing the basic standard of 80%, showing that the optimized treatment path and immersive environment significantly improve patient treatment engagement. The device operation convenience, function adaptability, and overall satisfaction ratings all exceeded 8.4 points, and the device discomfort incidence rate was only 5.3%, meeting the safety and comfort requirements for clinical application. In conclusion, the constructed VR-interactive mobile fusion immersive treatment environment has high user acceptance, providing important support for the clinical promotion and popularization of the technology.

6 CONCLUSION

This study focuses on the construction of an immersive music therapy environment combining VR and interactive mobile technology and optimization of psychological rehabilitation pathways. By designing a six-level collaborative architecture—comprising the Technology Layer, Perception Interaction Layer, Data Processing Layer, Algorithm Decision Layer, Application Layer, and Feedback

Optimization Layer—it achieved the modular construction of an immersive treatment environment and multimodal interaction. The study proposed a feature-layer-model-layer dual-fusion multimodal emotion recognition method, solving the challenges of accuracy and robustness in emotion recognition for patients with psychological disorders. A reinforcement learning-driven dynamic rehabilitation path optimization mechanism was built. With clinical experimental validation, the experimental group's depression and anxiety scale scores decreased by more than 35% from baseline, and the improvement in quality of life across various dimensions exceeded 27%, significantly outperforming the traditional treatment group. Technologically, the study has overcome three key issues: "synergy between immersive experience and mobile interaction," "fusion of multimodal data heterogeneity," and "personalized dynamic optimization of rehabilitation paths." Clinically, it offers a non-pharmacological intervention plan that is immersive, accurate, and personalized for patients with depression and anxiety disorders. This provides a new paradigm for the intelligent and digital development of psychological rehabilitation technologies.

7 REFERENCES

[1] Y. Deng *et al.*, "Global, regional and national burden of lung cancer attributable to $PM_{2.5}$ air pollution: Trends from 1990 to 2021 with projections to 2045," *Journal of Environmental Management*, vol. 390, p. 126216, 2025. https://doi.org/10.1016/j.jenvman.2025.126216

[2] M. Ahmad, N. Wahid, R. A. Hamid, S. Sadiq, and A. Mehmood, "Decision level fusion using hybrid classifier for mental disease classification," *Computers, Materials & Continua*, vol. 72, no. 3, pp. 5041–5058, 2022. https://doi.org/10.32604/cmc.2022.026077

[3] N. H. Shahimi, R. Lim, S. Mat, C. H. Goh, M. P. Tan, and E. Lim, "Association between mental illness and blood pressure variability: A systematic review," *Biomedical Engineering Online*, vol. 21, no. 1, p. 19, 2022. https://doi.org/10.1186/s12938-022-00985-w

[4] L. Liao, M. Du, and Z. Chen, "Air pollution, health care use and medical costs: Evidence from China," *Energy Economics*, vol. 95, p. 105132, 2021. https://doi.org/10.1016/j.eneco.2021.105132

[5] H. Alan, "A comprehensive evaluation of digital mental health literature: An integrative review and bibliometric analysis," *Behaviour & Information Technology*, vol. 44, no. 10, pp. 2282–2304, 2025. https://doi.org/10.1080/0144929X.2024.2303626

[6] S. Nepal *et al.*, "Capturing the college experience: A four-year mobile sensing study of mental health, resilience and behavior of college students during the pandemic," in *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 8, no. 1, 2024, pp. 1–37. https://doi.org/10.1145/3643501

[7] Q. Ding, "Evaluation of the efficacy of artificial neural network-based music therapy for depression," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, p. 9208607, 2022. https://doi.org/10.1155/2022/9208607

[8] M. Wang, G. Luo, and H. Chen, "Practice of music therapy for autistic children based on music data mining," *Mathematical Problems in Engineering*, vol. 2022, no. 1, p. 4576211, 2022. https://doi.org/10.1155/2022/4576211

[9] C. Y. Wu, "Music therapy music selection based on big data analysis," *Applied Mathematics and Nonlinear Sciences*, vol. 9, no. 1, pp. 1–12, 2024. https://doi.org/10.2478/amns-2024-0019

[10] M. Song and Y. M. Song, "Randomized controlled trials of digital mental health interventions on patients with schizophrenia spectrum disorder: A systematic review," *Telemedicine and e-Health*, vol. 29, no. 6, pp. 798–812, 2023. https://doi.org/10.1089/tmj.2022.0135

[11] C. Zhang, X. Wang, D. A. Juraev, R. F. Efendiev, and X.-G. Yue, "Risk research on blockchain technology in interactive mobile hospitals based on the entropy method," *International Journal of Interactive Mobile Technologies*, vol. 19, no. 10, pp. 152–162, 2025. https://doi.org/10.3991/ijim.v19i10.55483

[12] A. Dzedzickis, A. Kaklauskas, and V. Bucinskas, "Human emotion recognition: Review of sensors and methods," *Sensors*, vol. 20, no. 3, p. 592, 2020. https://doi.org/10.3390/s20030592

[13] B. Giardulli *et al.*, "Real and perceived feet orientation under fatiguing and non-fatiguing conditions in an immersive virtual reality environment," *Virtual Reality*, vol. 27, no. 3, pp. 2371–2381, 2023. https://doi.org/10.1007/s10055-023-00809-9

[14] S. Wei, "The role of EEG-based emotional feedback in enhancing the effectiveness of Tai Chi sports programs for oral health promotion," *IEEE Access*, vol. 13, pp. 81063–81082, 2025. https://doi.org/10.1109/ACCESS.2025.3561179

[15] H. Ren, L. Cheng, J. Zhang, and Q. Wang, "Eye-tracking investigation of emotional feedback to southern Hebei courtyard gates," *Journal of Asian Architecture and Building Engineering*, vol. 24, no. 6, pp. 5062–5079, 2024. https://doi.org/10.1080/13467581.2024.2407589

8 AUTHOR

Sai Wang holds a master's degree from Hebei University of Environmental Engineering. She is an assistant researcher, specializing in the history of music education, the historical development of urban music culture, and the study of modern and contemporary composers. She has led and participated in 8 provincial and municipal projects, published 10 papers, and edited 2 textbooks (E-mail: 18233545656@163.com).

Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 4 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i04.58587

PAPER

From Access to Achievement: A PLS-SEM Analysis of Mobile Learning Engagement in Chinese Higher Education

Qinghao Wu[1], Norhayati Mohd Yusof[2](✉), Zhijun Zhang[3]

[1]Shenyang Institute of Science and Technology, Liaoning, China

[2]Universiti Teknologi MARA (UiTM), Selangor, Malaysia

[3]Inner Mongolia Minzu University, Inner Mongolia, China

norhayatimy@uitm.edu.my

ABSTRACT

This study investigates the adoption, effectiveness, and pedagogical impact of mobile learning applications on the academic development of college students in China. Specifically, it examines how perceived ease of use, perceived usefulness, instructor support, and access to mobile technology influence academic outcomes, with mobile learning engagement acting as a mediating factor. A quantitative research design was employed using survey data collected from 234 college students across various Chinese universities. Validated scales were adopted from prior studies to measure all constructs. SmartPLS 4.0 was used to analyze the structural relationships through partial least squares structural equation modelling (PLS-SEM). Results confirmed that perceived ease of use, perceived usefulness, instructor support, and access to mobile technology all have significant positive effects on students' academic development. Furthermore, mobile learning engagement was found to mediate the relationships between these predictors and educational development, reinforcing its central role in digital learning environments. This study contributes to the growing body of literature on mobile learning by integrating elements from the technology acceptance model (TAM), Constructivist Learning Theory, and Engagement Theory. It offers practical and theoretical implications for educators, app developers, and policymakers aiming to foster meaningful student engagement and academic growth through mobile learning platforms.

KEYWORDS

mobile learning applications, academic development, mobile learning engagement, instructional support, technology adoption

1 INTRODUCTION

With the pace of the digital change gaining traction in our times, the adoption of mobile technologies in higher education has gained momentum, particularly in the example of China, where the national government has placed the development of

Wu, Q., Yusof, N. M., Zhang, Z. (2026). From Access to Achievement: A PLS-SEM Analysis of Mobile Learning Engagement in Chinese Higher Education. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(4), pp. 135–159. https://doi.org/10.3991/ijim.v20i04.58587

Article submitted 2025-09-09. Revision uploaded 2025-12-22. Final acceptance 2025-12-22.

innovative teaching and digital learning environments at the forefront of its agenda. The presence of mobile internet and smartphones has led to the view that mobile learning applications should be considered as an inexpensive and scalable method of supporting the learning process of students [1]. They offer diverse functions such as interactive course content, formative quizzes, discussion forums, and individualized learning paths [2]. A country like China has initiatives like the Smart Education of China project that has seen universities embrace mobile-based applications that not only have facilitated education inclusiveness but also pedagogical innovation with the aid of technology [3, 4]. Although such efforts have been made, the effectiveness of learning applications and their real usage still varies according to the region and university, which is why researchers seek to understand the psychological, technological, and environmental factors that influence the use of these applications by students [5]. In addition, Chinese higher education is facing increasing pressure to provide students with self-regulated learning abilities and digital literacy for the knowledge economy, so studying mobile learning environments is particularly pertinent [6].

Current empirical research brings strong evidence that factors influencing technology acceptance, specifically perceived ease of use and perceived usefulness, significantly contribute to students' intention to use educational apps as well as their learning results [7]. For example, several studies using the technology acceptance model (TAM) have shown that when students find a mobile app easy to use and helpful in studying, they are likely to utilize it regularly and gain academic advantages [9]. Within the Chinese context, studies have also established that perceived ease of use has a significant impact on the intention of students to use mobile learning platforms such as Super Star Learning, which are popularly utilized throughout Chinese universities [10]. Similarly, perceived usefulness has been related to higher levels of student satisfaction and higher motivation, and students have a higher probability of achieving higher academic outcomes, as students tend to internalize the relevance of these tools towards the achievement of their academic goal [11]. Combined, such findings support the idea that self-perceptions of the learning technologies among students are important factors to consider when defining their learning behaviors and academic performance.

Moreover, the studies have gradually highlighted the importance of contextual and social factors such as instructor support and access to mobile technologies in determining the learning process in students. One example is the support provided by an instructor, which has been found to raise students in terms of confidence in the usage of learning technologies, participation, and subsequent improvement in engagement and satisfaction [12]. In China, where teaching and learning practices are deeply rooted in Confucian values of teaching, the role of the instructor in facilitating and supporting the use of mobile learning tools in classroom activities also becomes more important [13]. In the meantime, the availability of mobile technology, both in terms of hardware and the ability to connect and be connected, and institutional infrastructures, are key factors in determining the success of learners in the digital learning setting. According to research carried out by [2, 14], students who have increased access to mobile phones and who have good Internet access demonstrate increased engagement and improved learning outcomes. Such empirical trends emphasize the multidimensionality of mobile learning and the need to have an integrative model that would take into consideration the perceptions of technology, teaching support, and infrastructure in the process of analyzing the learning growth.

Still, despite the growing amount of research, no one has managed to fill in some very important gaps, particularly in the Chinese educational background. To begin with, in as much as the TAM has been widely applied to explore the perceived ease of use and perceived usefulness, many studies have omitted these two variables as having direct impacts on the student performance without a close study of the mediating effects that transform technology perception to actual academic performance [15]. For instance, little has been studied regarding how mobile learning participation acts as a behavioral bridge between students' initial attitudes towards mobile applications and their academic achievements [16]. The lack of such mediation analysis restricts understanding why and how mobile apps can or cannot support learning achievements [17]. Additionally, there is less concern regarding the interplay between instructor guidance, technology access, and student engagement activities in mobile learning contexts, particularly in collectivist cultures such as China's, where social influence and teacher support are significant [18].

Secondly, the majority of existing studies in China tend to concentrate on individual platforms or assess disconnected variables, hence without a comprehensive framework that brings together user perceptions, contextual support, and behavioral engagement under one model. While research has investigated the effect of mobile learning on discrete skills like reading or the teaching of language [19], there is limited research that associates the more general construct of academic growth with these varied influences within a theoretically sound model. In addition, demographic and institutional variation in adopting mobile learning—e.g., urban vs. rural universities or first-rate vs. second-rate institutions—is usually not taken into account [20]. That limits the applicability and generalizability of delivered findings. Therefore, it is clear that integrative empirical research that will examine the direct association between the variables of technology and learning outcomes, as well as the mediating role of mobile learning involvement in Chinese higher education, is needed.

The study expects to fill gaps in existing empirical studies and enhance the knowledge on mobile learning in Chinese higher education by analyzing the interconnecting relationship between perceived ease of use, perceived usefulness, instructor support, mobile technology availability, mobile learning engagement, and academic growth of students. More specifically, the study will plan and analyze a conceptual framework that will unite cognitive, environmental, and behavioral variables within a mobile learning setting. The main focus is to understand not only whether these factors influence the process of academic development, but more importantly, how they influence it with the mediating role of student involvement in mobile learning technologies. With mobile learning reshaping the educational landscape, it is necessary to spend some time to go beyond the superficial analyses and examine how digital tools can influence learning behavior and learning outcomes.

Even though the study is framed within the framework of the Chinese higher education system, the issues that are discussed fall in a broader international discussion of mobile learning ecosystems as sources of digital equity and pedagogy-mediated technology-based innovation. According to such international standards as the ICT Competency Standards of the United Nations Educational, Scientific and Cultural Organization (UNESCO), the DigComp project by the European Commission, and the Digital Education Outlook published by the Organization of Economic Cooperation and Development (OECD), user-centered design of technologies, facilitation by educators, and meaningful engagement of learners are the priorities. These world plans underline the fact that the mobile learning uptake does not solely depend on the technological affordance, but it also depends upon a number of cultural, social, and institutional circumstances. Contextualizing the study in such a broader context will result in making it more pertinent since the Chinese case is not an exception of such a transnational issue on the degree to which higher education systems worldwide have adopted mobile learning to enhance the development of academic processes.

This study has an important theoretical and practical implication in the development of mobile learning in Chinese higher education. Theoretically speaking, the study adds to the currently existing discussion on educational technology by synthesizing the ideas of TAM, the theory of engagement, and contextual support frameworks into a single conceptual framework that discusses both direct and mediated pathways to academic development. It gives a more comprehensive insight into how mobile learning tools work in a rich learning system. Practically, the study can be useful to university administrators, instructional developers, and policymakers in identifying the most significant factors of student academic performance regarding mobile learning. Specifically, it underscores the necessity to ameliorate the technological component of the learning applications, as well as the support system and support measures surrounding its implementation. With the expanding dependence on digital platforms in post-pandemic China, the study provides valuable insights for developing inclusive, effective, and interactive mobile learning plans that are aligned with national educational priorities and institutional capabilities.

2 LITERATURE REVIEW

Mobile learning research has gathered pace throughout the world, with major contributions from Europe, South East Asia, North America, and Africa, which document similar determinants for the adoption of technology, such as student autonomy, teacher facilitation, and digital infrastructure. Studies conducted in Finland, Singapore, Canada, and South Korea emphasize that successful ecosystems for mobile learning depend on the combination of student-centered design, instructor scaffolding, and equitable access. Integrating these international studies offers a wider empirical base for the proposed framework and locates the current work in wider debates on digital learning readiness and technology-supported pedagogy in other parts of the world.

2.1 Perceived ease of use of learning apps and students' academic development

Perceived ease of use is the degree to which a person feels that applying a given technology will be effortless [20, 21]. In educational technologies, particularly learning applications, this aspect plays a pivotal role in influencing students' participation and inclination to adopt such tools as part of their academic practice [22]. Academic progress refers to the enhancement of knowledge gain, critical thinking, learning effectiveness, and general performance by students in learning activities [23, 24]. Learning applications that are developed with easy-to-use interfaces, simple navigation, and reduced cognitive load can promote more efficient learning experiences and lower the technological boundaries to educational participation [8]. The easier it is for students to use educational apps, the higher the chances they will check out their features, use them regularly, and incorporate them into their studying habits, which eventually leads to their academic advancement [6, 25]. Empirical findings have validated the relationship between perceived ease of use and educational achievement. [10, 26] found that ease of use significantly influences students' continued intention to use e-learning systems, which is associated with better learning outcomes. Similarly, it is reported that perceived ease of use positively correlates with satisfaction and learning performance among university students using learning platforms [27]. These results concur with the assumptions of the TAM, in which ease of use is a determining factor for behavior intention and system usage.

H1: Perceived ease of use of learning apps has a significant positive effect on students' academic development.

2.2 Perceived usefulness of learning apps and students' academic development

Perceived usefulness refers to the extent to which a user perceives that the use of a given system or tool will improve their work or task performance [28]. In schools, students believe that learning apps can enhance their academic achievement, learning effectiveness, or subject mastery. It is a dominant factor in determining whether or not students will use educational technology and use it regularly [29]. As students perceive that an electronic tool can offer helpful assistance in the realization of academic objectives, for example, through improved comprehension, interactive exercises, or prompt feedback, they will be more likely to use it actively [23, 30]. This way, perceived usefulness turns into a motivational construct that supports technology integration into routine academic conduct. Several empirical investigations verify the significant predictive contribution of perceived usefulness towards learning outcomes. Nuryakin et al. [31] discovered that the perceived usefulness of the students heavily determines their behavioral intention to use the learning technologies, which in turn impacts their learning performance. Furthermore, [32, 33] demonstrated that perceived usefulness had the highest influence among all factors in predicting the success of mobile learning systems within higher education. When learning applications are viewed as almost valuable for academic work, students tend to integrate them into their study patterns, thus improving their learning achievements.

H2: Perceived usefulness of learning apps has a significant positive effect on students' academic development.

2.3 Instructor support for mobile learning and students' academic development

Teacher support in mobile learning settings can be defined as the teaching, motivation, and pedagogical guidance that teachers provide to learners so that they can make use of mobile learning technologies in an effective manner [12]. This includes such elements as instruction scaffolds, appropriate communication of standards, regular feedback, and integration of mobile apps into formal teaching practice [18]. In this case, academic progress can be defined as the improvement of learning attainments, intellectual capacity, and scholarly drive of students [34]. The attitude of the students towards educational technology and their actual behavior is mainly influenced by the role of the instructors. When the instructors create a positive environment that encourages mobile learning and provide follow-up, students will feel encouraged and confident to apply learning applications to improve their academic competence [16, 35]. Empirical studies have highlighted the roles of teacher support to inform persuasive mobile learning experiences. It was determined that teacher engagement positively impacts the beliefs that students have about the usefulness and usability of mobile learning tools, which subsequently enhance the learning outcomes [13]. Similarly, [36] emphasized the fact that students working with the help of mobile learning platforms are more interested in using this kind of technology when teachers serve as a model and provide both technical and academic support. These findings show that professor support does not only facilitate technological adoption but also facilitates the learning value of mobile applications.

H3: Instructor support for mobile learning has a significant positive effect on students' academic development.

2.4 Access to mobile technology and students' academic development

The term access to mobile technology is used to denote the convenience and the practicability of accessing mobile phones, the internet, and applications of learning that are needed in mobile-assisted learning. It includes not only physical access (e.g., tablets, smartphones, data plans) but also digital access (e.g., access to apps, learning management systems, and learning material) [20]. With stable and equitable access, students will have an opportunity to participate in mobile learning processes, communicate with other students, and use digital resources to develop education [23, 37]. The potential of mobile learning cannot be fulfilled when there is inadequate access, particularly in a setting with scarce resources [14]. There exists empirical evidence that supports the claim that access to mobile technology is a starting-point facilitator of academic benefits of digital learning. As Sani and Ratri [38] point out, students who have uninterrupted access to the internet and the mobile gadgets are reported to be more engaged and satisfied in blended learning environments. In the same vein, Shortt et al. [29] demonstrated that the opportunity to have more access to mobile technologies correlates positively with the learning outcomes of students, particularly when accompanied by intelligently constructed instructional content. These findings point to the fact that access is not merely a logistical issue but a determinant of whether students can utilize mobile learning to the full to improve educational progress.

H4: Access to mobile technology has a significant positive effect on students' academic development.

2.5 Mobile learning engagement as mediator

Mobile learning engagement defines the level of student engagement, which is behavioral, emotional, and cognitive, in the active use of mobile-based applications and tools as learning ones [1]. This construct reflects several attributes of sustained attention, engagement, motivation, and determination of academic activities assisted with the help of mobile technologies by students. Although a perceived ease of use is the extent to which students hold that a learning app is easy and simple to use [2], its influence on academic advancement does not necessarily go directly.

Rather, it is the actual interaction with the tool that is performed as a mediating factor that directs this perception to a significant academic result. The previous research has shown that the usability is a significant factor that boosts the desire of students to use the digital tools, which consequently raises the engagement rates [5]. In addition, Budiarto et al. [6] have shown that as long as students consider mobile apps easy to use, they tend to be willing to develop positive emotional reactions and dedicate regular effort to academic interactions on the platform. On the same note, [10] discovered that the indirect effects of performance outcomes due to the perceived ease of educational technologies were by way of increased levels of cognitive and behavioral engagement. Since the processes of academic development are the processes of acquiring knowledge, critical thinking, and persistence in the work, one can suppose that it is only with the significant involvement of mobile learning tools that the ease of use can be transformed into academic benefits.

H5: Mobile learning engagement mediates the relationship between perceived ease of use of learning apps and students' academic development.

Although it is observed in previous literature that the perceived usefulness has a positive impact on the academic growth of students [34], more sophisticated studies indicate that the intensity of the mentioned association might be contingent upon the degree of learner engagement with the mobile platform. Mobile learning, the cognitive commitment, the interpersonal, and the interactive aspect of students in mobile-based educational activities can be the medium through which perceived usefulness can assert its pedagogical impact [18]. Studies conducted by Fan and Wang [22] indicate that those students who view mobile learning as an advantage will tend to spend more time engaging with course content, discussing, and taking tests on the mobile apps. Furthermore, this has been demonstrated by the works of [3, 39] that found that student engagement is a crucial mediator between technology-related perceptions and learning outcomes, and that usefulness perceptions cannot be effective on their own without active engagement. Therefore, the engagement of mobile learning may be theorized as the behavioral form of the student's confidence in the usefulness of an app that results in the transfer of abstract perception into concrete academic achievement.

H6: Mobile learning engagement mediates the relationship between perceived usefulness of learning apps and students' academic development.

Empirical studies have indicated that teacher support is a very important factor in improving the learning outcomes of students, as it encourages an environment where technology adoption and continued engagement become possible. According to [40], those students perceiving instructional support at high levels get more motivated to engage in mobile-based learning and demonstrate more persistence in using educational apps. Likewise, Saleem et al. also found that instructor guidance plays a central role in motivating and engaging students in mobile learning space [41]. Nevertheless, the direct influence of teacher support on academic growth might not be as complete a picture as it can be. Rather, the availability of instructor support can have an impact on the student outcomes by means of their interaction with the mobile learning in itself. With encouragement, feedback, and structured learning experiences delivered with the help of mobile apps, students are more susceptible to engaging in such applications, which results in more emotional and cognitive engagement [42]. It is important to note that [43] argue that such engagement is one of the key ways external supports determine academic outcomes. Additionally, Hazaymeh et al. [13] also showed that technology-related engagement behaviors mediated the role of instructional support on academic success significantly [36]. Based on these results, it is reasonable to presume that instructor support has a beneficial effect on the learning process of students, as it increases their interest in mobile learning.

H7: Mobile learning engagement mediates the relationship between instructor support for mobile learning and students' academic development.

Although access to mobile technology is an important facilitator of digital learning, previous studies indicate that the educational effect of access is not inherent and, in most cases, depends on the extent to which students utilize the technology that is made available [44]. Students who had increased access to mobile resources showed better academic performance [45]. Nevertheless, they had a strong influence on their interaction with mobile learning environments [46]. Similarly, [47] have found that access alone did not predict academic success, unless the students engaged in the tools provided in an active manner (taking quizzes, discussions, and learning the content) to do so. This is also in line with [10], who demonstrated that the usefulness of mobile learning was enhanced when students converted access to meaningful, repeated interaction. That is to say, the access enables the opportunity, though the student engagement is the one that brings to life the learning potential of mobile technologies [7]. It was also further highlighted [48] that the quality and frequency of interactions of students with mobile apps are the most critical predictors of their academic performance, despite the initial levels of access. These results indicate that mobile learning use is a very important mediator between access and education.

H8: Mobile learning engagement mediates the relationship between access to mobile technology and students' academic development.

2.6 Theoretical framework supporting the research

The theoretical perspectives over which the relationships in this study will be explained are mostly based on the TAM, as well as the contributions of Constructivist Learning Theory and the Engagement Theory. According to [49], TAM presupposes that there are two beliefs that clarify the intention of a person to use a technology that ultimately determines the actual use and the outcomes of it. This framework has been widely employed in education technology research in explaining the behavioral intention and learning outcomes among students in reaction to computer-mediated objects, including mobile learning applications [50]. The TAM is also used in the present study to provide a premise to assume that in case students find the learning apps very easy to use and useful, they will probably use them, which will contribute to the academic progress. Also, the study is informed by Constructivist Learning Theory, which centers on the active involvement of the learners and the creation of knowledge through interactions with learning environments [51]. As long as teachers support mobile learning applications and students have access to them, they allow self-directed and collaborative learning, therefore supporting the ideas of constructivism. Environmental enablers that facilitate this constructivist learning process are teacher support and access to mobile technology. The Engagement Theory [52] strengthens the model further by asserting that significant learning is achieved through engagement, especially when learners participate in interactive and meaningful activities. This research builds on existing frameworks for technology acceptance and mobile learning by incorporating behavioral engagement and contextual enablers into one model, namely instructor support and access to technology (see Figure 1). While TAM has traditionally focused on cognitive perceptions (ease of use and usefulness), it does not fully explain the behavioral mechanisms by which perceptions are translated into academic development. By proposing mobile learning engagement as a theoretically based mediator and embedding the contextual variables from constructivist and ecological learning perspectives, the study contributes to current knowledge and can provide a more holistic understanding of the working of mobile learning ecosystems in higher education settings.

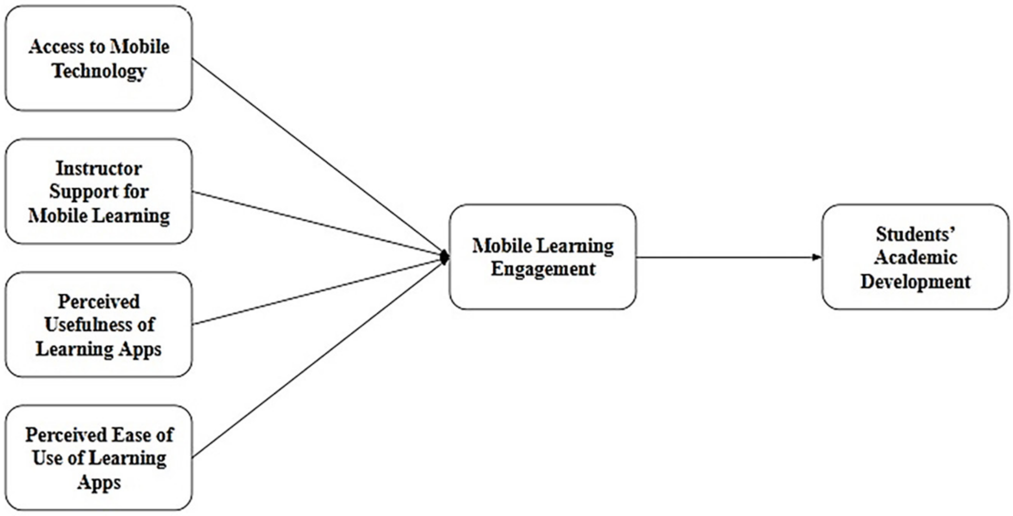


Fig. 1. Conceptual framework

3 METHODOLOGY

This study was conducted in a quantitative, cross-sectional survey design and was aimed at investigating the relationships between perceived ease of use, perceived usefulness, instructor support for mobile learning, access to mobile technology, mobile learning engagement, and academic development among undergraduate students in Chinese higher education. A survey approach was chosen due to the possibility of collecting standardized data from a relatively large student population engaged in mobile learning practices and is suitable for variance-oriented analytical techniques (partial least squares structural equation modelling—PLS-SEM).

3.1 Sampling and participants

The target population was undergraduate students in various universities in China if they frequently use mobile applications for academic purposes. A total of 234 valid responses were obtained by the convenience sampling technique. Although not very generalizable, convenience sampling can be widely applied in exploratory PLS-SEM studies because of its practicality and access to active users of technology. In order to have sufficient statistical power, sample size was assessed based on the "10-times rule," which states that the minimum sample should be 10 times the number of structural paths directed to the most complex endogenous construct. The construct "Academic Development" is assigned four structural paths, which require at least 40 cases. The sample size of 234 therefore surpasses this threshold and is suitable for the estimation of the models.

3.2 Instrument development

The questionnaire was divided into two sections: demographic and measurement items for the latent constructs. All constructs were measured through the use of established and validated scales adapted from previous studies. Items on perceived ease of use and perceived usefulness were adapted from Davis [49] and are widely used in studies of TAM. Instructor support and access to mobile technology scales were drawn from current mobile learning and educational technology literature. Because access to mobile technology can be seen as a broad enabling construct, the items included in the study intentionally span both physical access (e.g., availability of devices, stability of connectivity) and digital access (e.g., institutional access to platforms, availability of learning content). This approach is in agreement with research that recommends access be conceptualized as a multifaceted condition rather than a single technical indicator. Mobile learning engagement was assessed in terms of behavioral, emotional, and cognitive engagement indicators adapted from Fredricks et al. [53]. It is important to note that engagement was measured only based on perceptions. While self-reports are typical in mobile learning research, they represent perceptual rather than objective engagement behaviors. This is admitted to be a methodological limitation. All items used a five-point Likert scale varying from 1 ("strongly disagree") to 5 ("strongly agree"). The instrument was originally prepared in English and was translated into easy Chinese, and vice versa, to check for conceptual equivalence. A pilot test with 20 students proved clarity and reliability.

3.3 Data collection procedures

General mobile learning applications within the context of Chinese higher education would encompass institutional learning management application e.g. Chaoxing (Superstar Learning) and MOOC-based applications, discipline-specific education applications (e.g., language learning and test preparation applications), and communication and collaboration applications like WeChat that are often part of both formal and informal learning. The experience of students in Chinese higher education is characterized by a heterogeneous mobile learning environment, consisting of institutional learning management systems, discipline-specific educational applications, and communication tools, which are incorporated into the teaching and learning practices. Since students are usually able to move freely through these platforms in one course or learning task, confining the study to a particular application would have restricted the ecological validity of the engagement construct and reduced inference scope. Despite the lack of constraint on the use of a single platform, the items of the survey were tailored to reflect the overall interest of students in the use of mobile applications to achieve academic goals, where the Internet provided students with access to learning materials and interactive learning facilities and discussion with their teachers and lecturers. Data were gathered through an online survey that was distributed through university mailing lists, online learning communities, and popular Chinese platforms such as WeChat. Participation was voluntary and anonymous. The survey was open for four weeks so that there was time for students with various majors and universities to participate.

3.4 Ethical considerations

The study followed institutional and international ethics standards. Ethical approval was obtained from the Institutional Research Ethics Committee of the Beijing Normal University (Approval No. BNU-EDU-2024-117). Participants were provided with an informed consent statement that included information on the voluntary nature of participation, confidentiality of data, and the absence of personal identifiers. They were told of their right to withdraw at any time without penalty.

3.5 Data analysis

Data were analyzed with the software SmartPLS 4. The PLS-SEM procedure was based on a two-stage procedure. First, the measurement model was evaluated using indicator reliability, internal consistency reliability (Cronbach's alpha and composite reliability), convergent validity (average variance extracted), and discriminant validity (HTMT and Fornell-Larcker criteria). Following satisfactory results when measuring the measurement model, the structural model was evaluated. Path coefficients, t-values, and p-values were produced by a bootstrapping procedure with a subsample generation of the data (5000 subsamples). Mediation analysis was performed to investigate the indirect influence of perceived ease of use, perceived usefulness, instructor support, and access to mobile technology on academic development through mobile learning engagement. While statistical mediation assesses the importance of indirect effects in a numerical way, the study also used some theoretical justification. Drawing from Engagement Theory and Constructivist Learning Theory, engagement is conceptualized as the behavioral mechanism between perceptions and contextual supports and academic outcomes. As such, mediation is understood as a statistical effect as well as a theoretically based, explanatory process. Explanatory and predictive relevance (variance accounted for (VAF), R2, Q2, and model fit indices (SRMR, d_ULS)) were also examined.

4 RESULTS

Table 1 and Figure 2 show the reliability and validity measures for all the constructs employed within the study to ensure the soundness of the measurement model. Factor loadings (original sample estimate) for all the indicators are above the recommended 0.60, which ensures acceptable item reliability. All the items possess extremely significant t-values and p-values ($p < 0.001$), which reflect strong indicator reliability. Cronbach's Alpha coefficients for all variables are greater than 0.70, from 0.745 to 0.903, indicating good internal construct consistency. Composite reliabilities (CR) for all constructs are also significantly higher than 0.70, from 0.855 for perceived usefulness to 0.934 for perceived ease of use, supporting high construct reliability. In addition, the AVE values for all constructs surpass the suggested 0.50 cutoff value for convergent validity, confirming convergent validity. Overall, these findings suggest that scale items consistently and effectively measure their associated constructs and are amenable to structural equation modeling with SmartPLS.

Table 1. Variables' reliability and validity

Variables	Indicator	Original Sample	T Values	P Values	Cronbach's Alpha	CR	AVE
Academic Development	AD1	0.783	21.110	0.000	0.867	0.909	0.715
	AD2	0.842	29.442	0.000			
	AD3	0.906	56.711	0.000			
	AD4	0.848	32.668	0.000			
Access to Mobile Technology	AMT1	0.775	31.474	0.000	0.888	0.913	0.601
	AMT2	0.736	13.663	0.000			
	AMT3	0.811	32.629	0.000			
	AMT4	0.821	27.599	0.000			
	AMT5	0.626	9.401	0.000			
	AMT6	0.825	34.173	0.000			
	AMT7	0.812	23.526	0.000			
Instructor Support for Mobile Learning	ISML1	0.860	34.650	0.000	0.903	0.926	0.675
	ISML2	0.808	25.179	0.000			
	ISML3	0.857	38.971	0.000			
	ISML4	0.832	28.906	0.000			
	ISML5	0.828	25.976	0.000			
	ISML6	0.739	13.618	0.000			
Mobile Learning Engagement	MLE1	0.777	17.322	0.000	0.875	0.909	0.667
	MLE2	0.839	35.299	0.000			
	MLE3	0.845	32.580	0.000			
	MLE4	0.842	38.742	0.000			
	MLE5	0.779	16.557	0.000			
Perceived Ease of Use	PEOU1	0.928	58.939	0.000	0.895	0.934	0.826
	PEOU2	0.909	52.808	0.000			
	PEOU3	0.889	39.536	0.000			
Perceived Usefulness	PU1	0.809	23.711	0.000	0.745	0.855	0.662
	PU2	0.861	53.162	0.000			
	PU3	0.769	18.190	0.000			

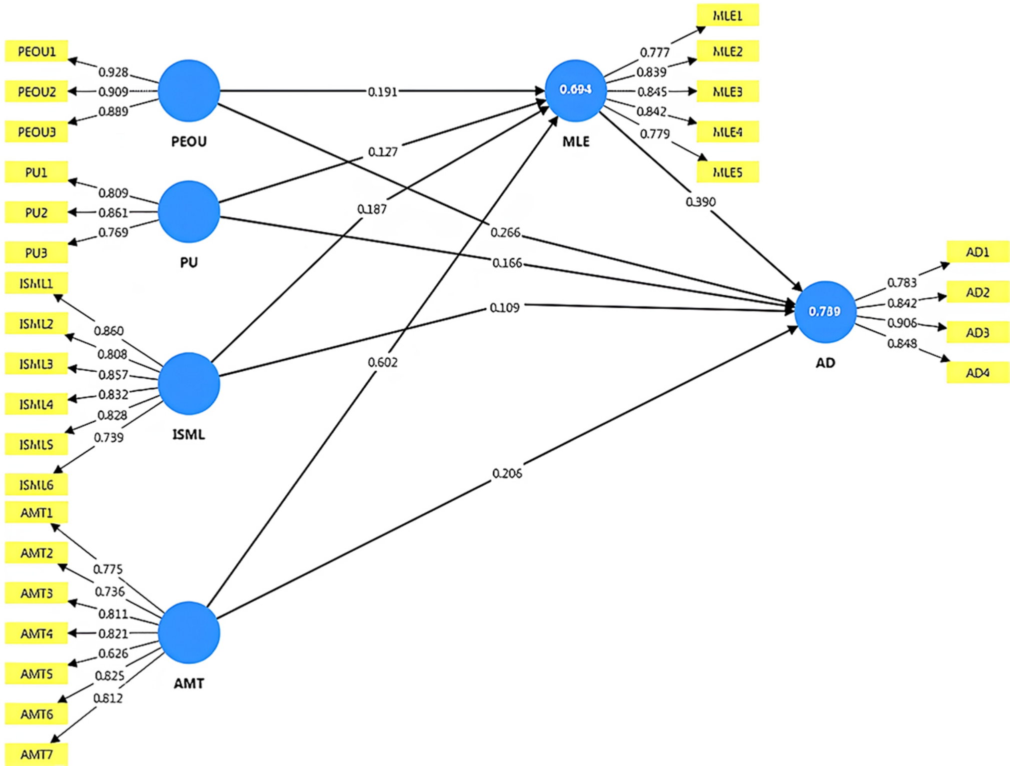


Fig. 2. Estimated model

Table 2 tests discriminant validity using the HMTM ratio and the Fornell-Larcker test. All the values in the HMT matrix fall below the 0.90 marker, which means that discriminant validity between the constructs is acceptable. Indicatively, the HMT of the mobile learning engagement and development of academic status is 0.868 with a high level of relationship, though the construct uniqueness is maintained. The Fornell-Larcker criterion restates discriminant validity by demonstrating that the square root of AVE (diagonal values) of each of the constructs is better than the correlations with other constructs. E.g., AVE has the square root of 0.909, which is greater than its correlation with other variables, such as access to mobile technology (0.534) and instructor support of mobile learning (0.540). This table provides confidence that all latent variables are empirically distinct from one another; that is, it confirms the design of the measurement model.

Table 2. Discriminant validity

	HTMT					
	AD	AMT	ISML	MLE	PEOU	PU
Academic Development						
Access to Mobile Technology	0.795					
Instructor Support for Mobile Learning	0.737	0.824				
Mobile Learning Engagement	0.868	0.834	0.865			
Perceived Ease of Use	0.776	0.588	0.599	0.648		
Perceived Usefulness	0.820	0.699	0.757	0.703	0.836	

(Continued)

Table 2. Discriminant validity *(Continued)*

	HTMT					
	AD	AMT	ISML	MLE	PEOU	PU
Fornell-Larcker criterion						
Academic Development	0.846					
Access to Mobile Technology	0.711	0.875				
Instructor Support for Mobile Learning	0.650	0.838	0.842			
Mobile Learning Engagement	0.759	0.839	0.771	0.847		
Perceived Ease of Use	0.697	0.534	0.540	0.582	0.909	
Perceived Usefulness	0.671	0.567	0.615	0.568	0.788	0.814

The explanatory power and effect size of the structural model are provided in Table 3. The R-Sq of the academic development is 0.739, which implies that about 74% of the academic development variance in students is attributed to the independent variables. Similarly, mobile learning engagement has a high percentage of explained variance, with an R-Sq of 0.694. These denote that the model is very predictive. The f-square results reveal that the engagement of mobile learning (0.130) and the availability of mobile technology (0.401) have medium effects on academic development. Such variables as perceived ease of use and perceived usefulness possess less significant effect sizes but still contribute to the model. The model fit indices also confirm the suitability of the model. The SRMR of the saturated model and the estimated model is 0.078, which is lower than the acceptable value of 0.08, which indicates a good fit. The structure fit of the model is also justified by the d ULS.

Table 3. R-square, f-square, and model fit statistics

	F Square		R Square	
	AD	MLE	R-Square	R-Square Adjusted
Students' Academic Development			0.739	0.736
Access to Mobile Technology	0.028	0.401		
Instructor Support for Mobile Learning	0.004	0.036		
Mobile Learning Engagement	0.130		0.694	0.690
Perceived Ease of Use of Learning Apps	0.081	0.051		
Perceived Usefulness of Learning Apps	0.030	0.002		
	MODEL FIT			
	Saturated Model		**Estimated Model**	
SRMR	0.078		0.078	
d_ULS	2.461		2.461	

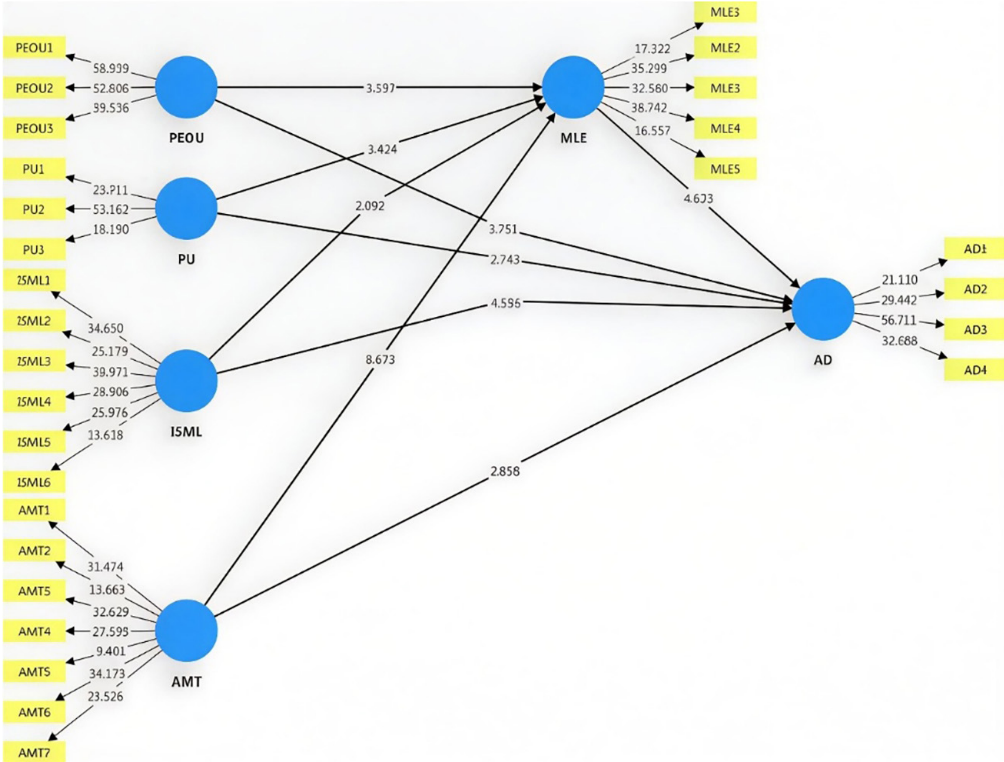


Fig. 3. Structural model for path analysis

Table 4 and Figure 3 indicate the results of path analysis, during which the direct and indirect relationship between the research variables is explored. The perceived ease of use has a significant direct correlation with academic development of 0.266 with a t-statistic of 3.751 and a p-value of 0.000. This implies that when the students find mobile learning apps to be user-friendly, it goes a long way into their academic growth. Similarly, there is a statistically significant and positive effect of perceived usefulness on academic development with a path coefficient of 0.166, a t-value of 2.743, and a p-value of 0.003. This implies that learners who consider mobile learning applications useful will be in a position to realize improvements in their academic results. The support of mobile learning by the instructor also shows a very high positive result of a 0.109 path coefficient, a t-value of 4.596, and a p-value of 0.000, which further reiterates the importance of the educator intervention in facilitating the academic growth of students. In addition, the mobile technology access has also shown a substantial influence on the academic development with a path coefficient of 0.206, a t-value of 2.858, and a p-value of 0.002. This finding indicates that the students who have a high level of access to mobile learning tools can probably perform well academically. Some of the key indirect impacts through mobile learning interaction are discovered in the analysis. The indirect relationship between the perceived ease of use and the academic development through mobile learning engagement is also significant with a path coefficient of 0.074, a t-value of 2.775, and a p-value of 0.003. This demonstrates that ease of use directly impacts academic performance and has an indirect impact on academic progress via engagement. Similarly, the indirect relationship of perceived usefulness on academic development through engagement has a significant path coefficient of 0.062 with a t-value of 2.480 and a p-value of 0.007, which means that perceived usefulness influences engagement and the opposite is also true. Moreover, mobile learning instructional support also has a strong indirect influence on academic development with a path coefficient of 0.073, a t-value of 1.798, and a p-value of 0.036. This highlights that instructional support triggers increased involvement in mobile learning that subsequently brings changes to academic development. The relevance of both the technological and pedagogical variables is justified by all the proposed relationships, which support the mediating effect of the mobile learning engagement on the academic performance of students.

Table 4. Path analysis

Hypotheses	Original Sample	Standard Deviation	T Statistics	P Values
PEOU –> AD	0.266	0.071	3.751	0.000
PU –> AD	0.166	0.061	2.743	0.003
ISML –> AD	0.109	0.024	4.596	0.000
AMT –> AD	0.206	0.072	2.858	0.002
PEOU –> MLE –> AD	0.074	0.027	2.775	0.003
PU –> MLE –> AD	0.062	0.025	2.480	0.007
ISML –> MLE –> AD	0.073	0.041	1.798	0.036

5 DISCUSSION

With the advent of mobile knowledge learning technologies, the higher education landscape was restructured and offered fresh avenues of academic engagement, personalized learning, and knowledge construction. Because of the increasing significance of digital literacy in Chinese higher education and the high rate of mobile device penetration, it is necessary to better understand how students interact with learning applications. The current study was intended to investigate the interconnection between the main constructs, such as the perceived ease of use, perceived usefulness, instructor support, access to mobile technology, and engagement in mobile learning, to influence the academic development of students. According to the TAM, Constructivist Learning Theory, and Engagement Theory, empirical evidence presented in this study can provide insights into the way in which the introduction of the use of digital tools in colleges contributes to the achievement of learning outcomes. The support for all hypothesized relationships highlights the dual role of technical and human aspects in promoting efficacious mobile learning experiences, underlining the central function of engagement as an intermediary mechanism.

The findings provide some actionable insights for international institutions of higher education. In systems with strong learner autonomy (e.g., Scandinavia, Canada), perceived usefulness may be the dominating driver of engagement. On the other hand, in teacher-centered cultures (e.g., East Asia, Middle East), instructor support may have a greater role in shaping learner behavior. Universities can thus make the model their own by emphasizing the contextual factor that is most compatible with their pedagogical norms and ensuring that engagement mechanisms stay front and center of mobile learning strategy.

The outcomes of this study offer strong evidence for Hypothesis 1, verifying that the ease of use perceived for learning apps has a substantial positive impact on students' academic progress. This is very much in line with the TAM, which holds that ease of use is a core driver of users' acceptance of technology in educational settings [49]. Suppose college students in China perceive learning applications as being intuitive and easy to use. In that case, they are likely to incorporate them into their daily academic tasks, which enhances their understanding, retention, and performance. Simple user interface, convenient access to academic materials, and little cognitive burden of using the functionality of applications will lead to increased utilization, which in turn will improve academic participation and advancement [54]. The obtained result can also be connected to the current literature in Chinese education settings, which emphasizes the role of student-centered environments in boosting the use of technology and the achievement of better learning outcomes. Moreover, based on the Constructivist Learning Theory, convenience improves active learning as it helps students to focus on knowledge creation and does not mess with technical problems.

Hypothesis 2 was also empirically validated, according to which the perceived usefulness of learning apps is a strong predictor of the academic developments of Chinese college students. TAM claims that the perceived usefulness directly influences user acceptance and the actual usage behavior [26]. When the students understand that a learning app will help them finish tasks more efficiently, they will better learn the course materials or even achieve higher academic results; therefore, there is a better status that they will use the technology. This observation is especially true in the case of the fast-growing education sector in China. The sufficiently high level of intelligent functionality in the majority of learning apps, such as customized quizzes, automated feedback, and real-time performance monitoring, produces the impression of academic progress, which results in the recurrence of use and more frequent interaction [6]. This connection is also promoted by Engagement Theory, as it observes that deep learning occurs when students are engaged in goal-focused activities using technology. Thus, not only does the utility of apps address an academic utilitarian need, but it also promotes higher-order cognition and self-directed learning.

The facts that support Hypothesis 3 show that the academic development of students positively depends on the support of mobile learning by their teachers. This fact highlights the importance of teachers in shaping the attitudes of students towards educational technology. Instructor support includes such activities as the recommendation of specific apps, technical support, the integration of app-related tasks into the classroom instruction, and the promotion of their frequent usage [1]. Such measures will make learners feel confident about using learning apps, which will boost their academic activities and performance. This is against Constructivist Learning Theory, which postulates that learning is a social process whereby teaching by an experienced teacher supplements cognitive construction. Also, the facilitator aspect of the instructor is complemented by the emphasis of Engagement Theory on real-world and participatory learning activities [53]. Teacher dominance practices in Chinese classrooms, where historically teacher dominance has been emphasized, may be highly influenced by instructor support of mobile learning technologies, and this will increase the effects of technology-enhanced learning.

The results are also favorable to Hypothesis 4, according to which the academic development of students is strongly affected by the access to mobile technology. Stable access to smartphones, tablets, and internet connectivity is one of the basic facilitators of mobile learning in the Chinese learning environment, where mobile phones are being used daily by young people [14]. Students who have consistent access to mobile phones will have a greater chance of using learning applications to learn, consolidate knowledge, collaborate with others, and manage academic assignments. This is in line with studies that have established the relevance of digital access and access to infrastructure in the success of mobile learning [19]. The access allows the students to customize their learning conditions, and students are now able to study at any time and location, thus extending the learning past the classroom boundaries. Engagement Theory supports this finding by saying that the technological situation must be attainable and consistent to ensure that learners can successfully use the online materials and socialize with their classmates. In this way, the institutional and governmental measures aimed at reducing the digital divide as well as enhancing access to mobile learning facilities in China are fundamental in ensuring that technology-enhanced education maximum benefits are realized.

The fact that Hypothesis 5 gets accepted means that learning apps can be perceived as easy to use by students, and such a perception encourages them to engage in the mobile learning process to an even greater degree, which in its turn leads to improved academic development. Similarly, the adoption of Hypothesis 6 would suggest that the perceived usefulness of learning apps would lead to increased learning engagement that further leads to better academics. Such mediated relationships have strong foundations on TAM, in which perceived ease of use and perceived usefulness are observed to be important antecedents of behavioral intention to use technology, which leads to the eventual patterns and consequences of the use [49]. The intervention of mobile learning interaction introduces a deeper set of how these perceptions are being carried into academic performance. Constructivist Learning Theory provides the description of the process by the following statement: valuable learning happens when students are engaged with the material in a manner of autonomous interaction. Reflective, self-paced, and individualized learning can be achieved with the help of learning applications that are practical and easy to navigate. The explanation is complemented by the Engagement Theory [52], which states that authentic learning occurs when learners are engaged both mentally and behaviorally in interactive processes. Therefore, mobile engagement is the interface between the positive attitude towards technology and the actual academic performance of the students, which gives a theoretical and empirical basis to the development of the student-centered digital learning spaces.

The Hypothesis 7 that was accepted presupposes that teacher support appropriate to mobile learning, including the use of apps in classes, educating students on app usage, and encouraging them to learn more about educational technologies, yields a better apprehension of students' engagement. It, hence, results in high academic performances. On the same note, Hypothesis 8 evidence also states that access to mobile technology is not sufficient, but rather, it is the interaction with that technology that facilitates access to translate into valid learning outcomes [55, 56]. These findings support the fundamental principles of Constructivist Learning Theory that postulate facilitation of scaffolding by the teacher and the active role that the learner plays towards developing knowledge based on the interactions between the learner and the environment. This is supported by Engagement Theory, which opines that the access of the interactive and collaborative digital resources should be complemented with valuable engagement to result in academic achievement [57, 58]. Mobile learning engagement as a psychological and behavioral process through mediating the relationship between contextual support (instructor support and access) and academic growth represents a core psychological and behavioral process where the enabling conditions are transformed into measurable academic outcomes [38]. These discoveries point to the necessity of providing Chinese academic institutions with not only access and teacher training but also the development of learning strategies that can ensure long-term involvement of the students in using the mobile learning means.

In general, the findings of the current research are a detailed description of the effect of mobile learning applications on academic success through an intricate set of perceptions, resources, and student habits. All the hypotheses under support show that involvement of students is a significant middle factor between the acceptance of technology and academic performance, and that the adoption of mobile tools is not sufficient without active and meaningful use. The integration of TAM and constructivist and engagement models has remained one of the robust theoretical perspectives in which to understand the Chinese academic mobile learning dynamics. The results are both theoretical and practical for the educationists, app developers, and policymakers in the quest to enhance student achievement in digital innovation. Lastly, the study justifies that when the students are assisted, feel they are getting something of worth, and are inspired to get deeply involved with technology, their schooling career might significantly increase.

6 CONCLUSION

Overall, the study offers information on the ideal significance of mobile learning applications in the establishment of the learning outcomes among Chinese college students. By reviewing the impact of perceived ease of use, perceived usefulness, pedagogic support, and availability of mobile technology on learning and the mediating role of mobile learning engagement, the work gives a comprehensive vision of how digital resources can influence learning outcomes. The ease of all the hypotheses confirms the superiority of user perception, environmental support, and digital accessibility in improving student performance through mobile platforms. Following the TAM with the contribution of the Constructivist Learning Theory and the Engagement Theory, the findings indicate not only the technological but also the pedagogical and engagement-based elements of digital learning. The study confirms that the promotion of active involvement with the help of mobile devices can bridge the divide between the use of technologies and significant academic success. Overall, the study has both theoretical and practical implications, as it introduces a sophisticated framework that can be used by educators, developers, and policymakers to develop more efficient mobile learning platforms that can support the evolving demands of modern students in institutions of higher learning.

7 IMPLICATIONS

7.1 Practical implications

Although the findings are considered within the Chinese higher education framework, they can be transferred to the learning abroad mobile learning programs. The identified core mechanisms, the perception of technology by students, and the support of the instructors with the context, the access to the mobile devices, and the participation are regarded as typical factors of the digital learning frameworks worldwide. Nonetheless, their impact may vary in relation to the cultural norms, institutional roles, and levels of digital infrastructure maturity. Perceived usefulness should mediate engagement in situations where the degree of autonomy of the learner is high and where learning is decentralized, whereas in situations where instructor power is important in the determination of learning behavior, pedagogical support can play a bigger mediating role. The applicability of the research thus is not limited to China, as long as its application considers contextual factors like the digital equity, pedagogical culture, and the institutional support systems. This cross-cultural awareness is contributing to the value of this study to the global discourse on mobile learning ecosystems.

The findings of this study offer meaningful practical implications to educators, institutional policymakers, and ed-tech developers wishing to enhance the academic growth of college students with the help of mobile learning. First of all, the fact that the perceived ease of use and perceived usefulness were on the forefront of academic development suggests that Chinese institutions should invest in a user-friendly and pedagogically significant mobile app that has been integrated near the academic goals of the students. Teachers must be motivated to actively use such applications in their teaching practice and provide continuous support, which will help students gain confidence and motivation. Moreover, by enabling more students to access mobile technology, either by institutionalizing it or by upgrading their learning infrastructure, they can significantly increase their learning. Training of the faculty and students on how to use educational apps productively could also enhance the learning process. The developers of the apps should also incorporate interactive, personalized, and engagement-based features in their strategy because student engagement has become one of the determinants of their academic output. Combining user-friendly design, content, and teaching aids, the stakeholders are able to create a mobile learning environment that is not merely technologically on the cutting edge but also effective in terms of academics.

7.2 Theoretical implications

Theoretically, this study makes an addition to the body of knowledge on technology acceptance and academic development because it develops the TAM by additions of constructs of Constructivist Learning Theory and Engagement Theory. By the depiction of the mediating role of mobile learning engagement, the study supplements the classical aspect of TAM in its focus on behavioral intention and system use by a pedagogically meaningful outcome: academic development. This extension offers a more holistic approach to assessing technology in educational settings, especially in the fast-changing world of mobile learning. In addition, by confirming the role of external variables such as teacher support and facility access to mobile technology, the study highlights contextual and environmental enablers that are frequently underemphasized in technology adoption frameworks. The convergence of constructivist and engagement theories also emphasizes learners' active construction of knowledge and improvement of academic outcomes through interactive digital encounters. This study, therefore, provides the framework for future theoretical frameworks that are technology-based but pedagogically enriched, particularly in Asian higher education contexts. This study is an improvement of the classical TAM, as it entails contextual and pedagogical aspects that enhance its theoretical context. Traditional TAM applications mainly focus on the perceptions of ease of use and usefulness of the users as the factors influencing behavioral intention and use of the system. Nonetheless, the study adds more elements to the model by incorporating instructor support and access to mobile technology as important contextual variables based on the socio-pedagogical environment where learning technologies are implemented.

Moreover, the incorporation of mobile learning engagement as a mediating variable takes TAM to a new stage of being a predictive model of technology acceptance to an explanatory model of academic growth. This theoretical extension recognizes that the process of technology acceptance in education is not an entirely cognitive process but rather is influenced by instructional context, accessibility, and long-term behavioral engagement. Therefore, the study has a role to play in the development of debates about the flexibility of TAM to education systems and its applicability in examining digital learning practices in different learning institutions.

7.3 Limitations and future directions

The current study has several limitations. The use of convenience sampling limits the generalizability of the results to larger populations of students. Although the study results are robust and internally consistent, it has to be admitted that there are several limitations. The convenience sampling and sample size of 234 undergraduate students, despite being adequate to estimate PLS-SEM and adhering to the ten times rule, limit the generalizability of the findings. The system of higher education in China is very heterogeneous with respect to institutional levels, its characteristics in the development of different regions, and digital infrastructure, and the sample under consideration might be restricted to represent this variety in all its aspects. The results can therefore be taken as a guide and not a reflection of all the Chinese universities. All constructs were assessed based on the self-reported perceptions, which include mobile learning engagement, which may not necessarily reflect actual behavioral patterns and actual test scores due to social desirability or recall biases. The lack of objective behavioral data (such as app usage logs) is a limitation with the engagement construct, as it compromises the precision of engagement. The cross-sectional design rules out any causal inferences, while contextual variables such as digital literacy, socioeconomic background, and discipline-specific learning environments were not controlled. Future research should include mixed-method methodologies, behavioral analytics, and sampling strategies that are more diverse in order to increase explanatory depth and external validity. In future studies, more extensive and multi-site big sampling techniques, such as stratified or probability sampling techniques, must be used to provide better external validity and provide stronger conclusions about populations. Longitudinal and mixed-method designs might help supplement the validity of the proposed model in the context of various institutional and regional environments.

Future research can further the research done in this work by looking at the longitudinal designs in order to gain more knowledge into the causal relationships and the long-term effects of mobile learning in academic development. The sample of students representing different regions, institutions, and students of different levels should contribute to the external validity of the results. The inclusion of moderating or control variables such as gender, discipline in school, or technology experience was also a possibility that researchers could use to unlock further insights. In addition to quantitative surveys, the qualitative data collection tools, i.e., interviews or focus groups, may also be used to learn more about the mobile learning engagement. Further studies on the effectiveness of the diverse mobile learning applications, such as gamified applications, collaborative applications, and discipline-specific applications, can be conducted to establish the most effective application that can lead to better learning outcomes in different learning environments.

8 REFERENCES

[1] N. A. A. Abdelwahed and B. A. Soomro, "Attitudes and intentions towards the adoption of mobile learning during COVID-19: Building an exciting career through vocational education," *Education + Training*, vol. 65, no. 2, pp. 210–231, 2023. https://doi.org/10.1108/ET-02-2022-0048

[2] A. Alam and A. Mohanty, "Learning on the move: A pedagogical framework for state-of-the-art mobile learning," in *Proc. Int. Conf. Data Management, Analytics & Innovation*, 2023. https://doi.org/10.1007/978-981-99-1414-2_52

[3] C.-C. Chen, C.-C. Liu, T.-H. Chiu, Y.-W. Lee, and K.-C. Wu, "Role of perceived ease of use for augmented reality app designed to help children navigate smart libraries," *Int. J. Human-Computer Interaction*, vol. 39, no. 13, pp. 2606–2623, 2023. https://doi.org/10.1080/10447318.2022.2082017

[4] S. Papadakis, S. O. Semerikov, and A. M. Striuk, "Embracing digital innovation and cloud technologies for transformative learning experiences," in *Proc. 11th Workshop Cloud Technologies in Education (CTE 2023)*, vol. 3679, 2023, pp. 1–21.

[5] S. A. Booton, A. Hodgkiss, and V. A. Murphy, "The impact of mobile application features on children's language and literacy learning: A systematic review," *Computer Assisted Language Learning*, vol. 36, no. 3, pp. 400–429, 2023. https://doi.org/10.1080/09588221.2021.1930057

[6] M. K. Budiarto, G. Gunarhadi, and A. Rahman, "Technology in education through mobile learning application (MLA) and its impact on learning outcomes: Literature review," *J. Education and Learning (EduLearn)*, vol. 18, no. 2, pp. 413–420, 2024. https://doi.org/10.11591/edulearn.v18i2.20976

[7] M. Al-Emran and C. Griffy-Brown, "The role of technology adoption in sustainable development: Overview, opportunities, challenges, and future research agendas," *Technology in Society*, vol. 73, p. 102240, 2023. https://doi.org/10.1016/j.techsoc.2023.102240

[8] K. Çelik and A. Ayaz, "Evaluation of metaverse use intention in software education of university students: Combining technology acceptance model with external variables," *Educational Technology Research and Development*, vol. 73, no. 1, pp. 641–662, 2025. https://doi.org/10.1007/s11423-024-10415-4

[9] J. Díaz-Arancibia *et al.*, "Navigating digital transformation and technology adoption: A literature review from small and medium-sized enterprises in developing countries," *Sustainability*, vol. 16, no. 14, p. 5946, 2024. https://doi.org/10.3390/su16145946

[10] B. Chung and N. Lo, "A comparative study of self-regulated English learning through mobile language-learning applications in post-pandemic Hong Kong and South Korea," *Smart Learning Environments*, vol. 12, no. 1, pp. 1–18, 2025. https://doi.org/10.1186/s40561-025-00399-w

[11] K. Lalani, J. Crawford, and K. Butler-Henderson, "Academic leadership during COVID-19 in higher education: Technology adoption and adaptation for online learning during a pandemic," *Int. J. Leadership in Education*, vol. 28, no. 1, pp. 1–17, 2025. https://doi.org/10.1080/13603124.2021.1988716

[12] V. P. Dennen and M. K. Jones, "The role of the online instructor: A nexus of skills, activities, and values that support learning," in *Handbook of Open, Distance and Digital Education*, Springer, 2023, pp. 1073–1088. https://doi.org/10.1007/978-981-19-2080-6_62

[13] W. A. Hazaymeh, A. Bouzenoun, and A. Remache, "EFL instructors' perspective on using AI applications in English as a foreign language teaching and learning," *Emerging Science Journal*, vol. 8, no. 5, pp. 73–87, 2024. https://doi.org/10.28991/ESJ-2024-SIED1-05

[14] I. Nicolaidou, P. Pissas, and D. Boglou, "Comparing immersive virtual reality to mobile applications in foreign language learning in higher education: A quasi-experiment," *Interactive Learning Environments*, vol. 31, no. 4, pp. 2001–2015, 2023. https://doi.org/10.1080/10494820.2020.1870504

[15] X. Wang and B. L. Reynolds, "Beyond the books: Exploring factors shaping Chinese English learners' engagement with large language models for vocabulary learning," *Education Sciences*, vol. 14, no. 5, p. 496, 2024. https://doi.org/10.3390/educsci14050496

[16] S. Attuquayefio, D. Aboagye-Darko, and A. Q. Okronipa, "Digital academic entrepreneurship in emerging economies: Antecedents of social media adoption for academic entrepreneurship," *Education and Information Technologies*, vol. 29, no. 10, pp. 11765–11791, 2024. https://doi.org/10.1007/s10639-023-12286-3

[17] R. Lopez-Chila, N. Mora-Chiquito, and D. Saavedra-Muñoz, "Enhancing the student academic profile: Trend analysis of technology adoption in automotive engineering," in *Proc. 2025 IEEE Engineering Education World Conf. (EDUNINE)*, 2025. https://doi.org/10.1109/EDUNINE62377.2025.10981359

[18] J. N. Lyanda, G. A. Koteng, and R. O. Ong'unya, "School administration support systems for educational technology adoption and students' academic achievement in secondary schools in Kenya," *African Journal of Empirical Research*, vol. 4, no. 2, pp. 363–374, 2023. https://doi.org/10.51867/ajernet.4.2.36

[19] M. Z. M. Fuzi and W. A. J. W. Yahaya, "Empowering problem-solving in computer science: A need analysis for a computational thinking mobile learning application," *Asian Journal of Research in Education and Social Sciences*, vol. 6, no. 1, pp. 408–417, 2024.

[20] J. Khlaisang, N. Songkram, F. Huang, and T. Teo, "Teachers' perception of the use of mobile technologies with smart applications to enhance students' thinking skills," *Interactive Learning Environments*, vol. 31, no. 8, pp. 5037–5058, 2023. https://doi.org/10.1080/10494820.2021.1993933

[21] A. Woick, H. Rinn, L. Grogorick, T. Mühleisen, and D. Markgraf, "Metaverse in higher education – A systematic literature review," in *37th Bled eConference Resilience Through Digital Innovation: Enabling the Twin Transition*, 2024. https://opus4.kobv.de/opus4-UBICO/frontdoor/index/index/year/2024/docId/33665

[22] J. Fan and Z. Wang, "The impact of gamified interaction on mobile learning app users' learning performance," *Behaviour & Information Technology*, vol. 44, no. 7, pp. 1306–1319, 2025. https://doi.org/10.1080/0144929X.2020.1787516

[23] S. Papadakis *et al.*, "Advancing lifelong learning with AI-enhanced ICT," in *Proc. IX Int. Workshop on Professional Retraining and Life-Long Learning using ICT (3L-Person 2024)*, 2024, pp. 1–9.

[24] H. Wu, Y. Wang, and Y. Wang, "To use or not to use? Determinants of EFL learners' behavioral intention to use AI," *Int. Rev. Research in Open and Distributed Learning*, vol. 25, no. 3, pp. 158–178, 2024. https://doi.org/10.19173/irrodl.v25i3.7708

[25] S. Bedenlier *et al.*, "Facilitating student engagement through educational technology in higher education," *Australasian Journal of Educational Technology*, vol. 36, pp. 126–150, 2020. https://doi.org/10.14742/ajet.5477

[26] K. Daniel *et al.*, "Motivate students for better academic achievement," *Computer Applications in Engineering Education*, vol. 32, no. 4, p. e22733, 2024. https://doi.org/10.1002/cae.22733

[27] D. Y. Park and H. Kim, "Determinants of intentions to use digital mental healthcare content," *Sustainability*, vol. 15, no. 1, p. 872, 2023. https://doi.org/10.3390/su15010872

[28] N. Yao and Q. Wang, "Factors influencing pre-service special education teachers' intention toward AI in education," *Heliyon*, vol. 10, no. 14, 2024. https://doi.org/10.1016/j.heliyon.2024.e34894

[29] M. Shortt *et al.*, "Gamification in mobile-assisted language learning," *Computer Assisted Language Learning*, vol. 36, no. 3, pp. 517–554, 2023. https://doi.org/10.1080/09588221.2021.1933540

[30] A. A. Linus *et al.*, "Perceived usefulness, ease of use, and intention to utilize online tools for learning," *Indonesian Journal of Multidisciplinary Research*, vol. 5, no. 1, pp. 41–52, 2025. https://doi.org/10.17509/ijomr.v5i1.81387

[31] N. Nuryakin, N. L. P. Rakotoarizaka, and H. G. Musa, "The effect of perceived usefulness and perceived ease of use on student satisfaction," *APMBA*, vol. 11, no. 3, pp. 323–336, 2023. https://doi.org/10.21776/ub.apmba.2023.011.03.5

[32] Z. Luo, "Determinants of the perceived usefulness in gamification for ESL teaching," *Education and Information Technologies*, vol. 28, no. 4, pp. 4741–4768, 2023. https://doi.org/10.1007/s10639-022-11409-6

[33] R. Scherer *et al.*, "Profiling teachers' readiness for online teaching," *Computers in Human Behavior*, vol. 118, p. 106675, 2020. https://doi.org/10.1016/j.chb.2020.106675

[34] Z. Meng and R. Li, "Understanding Chinese teachers' informal online learning continuance," *Journal of Computing in Higher Education*, vol. 36, no. 2, pp. 275–297, 2024. https://doi.org/10.1007/s12528-023-09352-7

[35] M. Matsieli and S. Mutula, "COVID-19 and digital transformation in higher education institutions," *Education Sciences*, vol. 14, no. 8, p. 819, 2024. https://doi.org/10.3390/educsci14080819

[36] J. Xie and A. P. Correia, "The effects of instructor participation in asynchronous online discussions," *British Journal of Educational Technology*, vol. 55, no. 1, pp. 71–89, 2024. https://doi.org/10.1111/bjet.13350

[37] B. Kristanto, T. Glomjai, and D. Putri, "Enhancing nursing students' long-term retention through multimedia mobile learning," *Journal of Advanced Health Informatics Research*, vol. 2, no. 1, pp. 12–23, 2024. https://doi.org/10.59247/jahir.v2i1.178

[38] J. Sani and D. Ratri, "Effectiveness of gamification elements in mobile-based English learning," *International Journal of Language Education and Cultural Review*, vol. 10, no. 2, pp. 213–226, 2024. https://doi.org/10.21009/ijlecr.v10i2.49483

[39] H. M. Iddrisu, S. A. Iddrisu, and B. Aminu, "Gender differences in the adoption of AI writing tools," *International Journal of Educational Innovation and Research*, vol. 4, no. 1, pp. 110–111, 2025. https://doi.org/10.31949/ijeir.v4i1.11717

[40] T. Wang *et al.*, "Exploring the role of AI assistants in computer science education," in *Proc. IEEE Symp. Visual Languages and Human-Centric Computing (VL/HCC)*, 2023. https://doi.org/10.1109/VL-HCC57772.2023.00018

[41] F. Saleem, E. Chikhaoui, and M. I. Malik, "Technostress in students and quality of online learning," *Frontiers in Education*, vol. 9, 2024. https://doi.org/10.3389/feduc.2024.1309642

[42] X. Liu and X. Shao, "Modern mobile learning technologies in online piano education," *Interactive Learning Environments*, vol. 32, no. 4, pp. 1279–1290, 2024. https://doi.org/10.1080/10494820.2022.2118787

[43] J. Garzón and G. Lampropoulos, "Mobile learning for science education," *Interactive Learning Environments*, vol. 32, no. 10, pp. 6735–6750, 2024. https://doi.org/10.1080/10494820.2023.2280973

[44] R. Zhi, Y. Wang, and Y. Wang, "The role of emotional intelligence and self-efficacy in EFL teachers' technology adoption," *Asia-Pacific Education Researcher*, vol. 33, no. 4, pp. 845–856, 2024. https://doi.org/10.1007/s40299-023-00782-6

[45] M. M. M. Fuzi and W. A. J. W. Yahaya, "Assessing the role of mobile applications in improving stoichiometry understanding," *Int. J. Advanced Research in Education and Society*, vol. 6, no. 1, pp. 349–360, 2024.

[46] M. Mannan *et al.*, "Technology adoption for higher education in Bangladesh," *Journal of Education and Social Sciences*, vol. 24, no. 1, pp. 1–9, 2023.

[47] L. Zhou and G. M. Alam, "Commercial higher education strategies for recruiting international students in China," *Discover Sustainability*, vol. 5, no. 1, p. 33, 2024. https://doi.org/10.1007/s43621-024-00216-3

[48] P. Sun *et al*., "Investigating students' behavioral intention to use ChatGPT for educational purposes," *Sustainable Futures*, vol. 9, p. 100531, 2025. https://doi.org/10.1016/j.sftr.2025.100531
[49] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, vol. 13, no. 3, pp. 319–340, 1989. https://doi.org/10.2307/249008
[50] T. Teo, "Factors influencing teachers' intention to use technology," *Computers & Education*, vol. 57, no. 4, pp. 2432–2440, 2011. https://doi.org/10.1016/j.compedu.2011.06.008
[51] D. Laurillard, *Teaching as a Design Science*. New York, NY: Routledge, 2013. https://doi.org/10.4324/9780203125083
[52] S. Marshall, "Engagement theory, WebCT, and academic writing in Australia," *Int. J. Education and Development using ICT*, vol. 3, no. 2, pp. 109–115, 2007. http://ijedict.dec.uwi.edu/viewarticle.php?id=227
[53] J. A. Fredricks, M. Filsecker, and M. A. Lawson, "Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues," *Learn. Instr*., vol. 43, pp. 1–4, 2016. https://doi.org/10.1016/j.learninstruc.2016.02.002
[54] V. Chan, "Impact of an extended reality-powered mobile application," *Interactive Learning Environments*, pp. 1–23, 2025. https://doi.org/10.1080/10494820.2025.2519702
[55] C. A. Eden, O. N. Chisom, and I. S. Adeniyi, "Online learning and community engagement," *World Journal of Advanced Research and Reviews*, vol. 21, no. 3, pp. 232–239, 2024. https://doi.org/10.30574/wjarr.2024.21.3.0693
[56] J. Zhao and N. Zhao, "The impact of interactive mobile learning on enhancing university students' English-speaking proficiency," *Int. J. Interact. Mobile Technol. (iJIM)*, vol. 18, no. 24, pp. 130–144, 2024. https://doi.org/10.3991/ijim.v18i24.53093
[57] W. Zhao, "Driving the integration of mobile learning and blended learning models in higher education," *Int. J. Interact. Mobile Technol. (iJIM)*, vol. 19, no. 5, pp. 45–59, 2025. https://doi.org/10.3991/ijim.v19i05.54529
[58] H. Li, K. Numtong, D. Gan, and W. P. Ngern, "Mapping mobile learning adoption in online education: A BERTopic review of TAM studies (2020–2024)," *Int. J. Interact. Mobile Technol. (iJIM)*, vol. 19, no. 19, pp. 19–38, 2025. https://doi.org/10.3991/ijim.v19i19.56907

9 AUTHORS

Qinghao Wu is a Lecturer at the College of Humanities, Shenyang Institute of Science and Technology, 110166, Liaoning, China. She is currently pursuing her Ph.D. in Education at the Faculty of Education, Universiti Teknologi MARA (UiTM), Malaysia. Research interests lie in the fields of Mobile Learning, Educational Technology, E-learning, and Blended Learning (E-mail: wqh.1214@163.com).

Dr. Norhayati Mohd Yusof is a Senior Lecturer at the Faculty of Education, Universiti Teknologi MARA (UiTM), 42300 Selangor, Malaysia. She holds a PhD in Education specializing in Instructional Technology. Her teaching and supervision focus on educational technology, digital pedagogy, qualitative research methodology, and technology-enhanced learning. Her research interests include digital learning environments, metaverse and avatar-based learning, e-personality development, netnography, and self-regulated digital engagement in education (E-mail: norhayatimy@uitm.edu.my).

Zhijun Zhang is a researcher at the Office of Academic Affairs, School of Foreign Languages, Inner Mongolia Minzu University, 028000, Inner Mongolia, China. Research interests in the fields of course evaluation and e-learning (E-mail: 183658922@qq.com).

Imprint

iJIM – International Journal of Interactive Mobile Technologies
http://www.i-jim.org

Indexing
International Journal of Interactive Mobile Technologies is indexed in
Elsevier Scopus, Ulrich, EBSCO, Google Scholar, and DBLP.

Publication Frequency
Bimonthly

Publisher
International Federation of Engineering Education Societies (IFEES)
IFEES-GEDC Secretariat
c/o George Mason University
Volgenau School of Engineering
4400 University Drive, MS 4A3
Fairfax, VA 22030
USA

Editorial Office
CTI Global
Frankfurt, Vienna, New York, Bengaluru, Hong Kong
Madrider Strasse 4
60327 Frankfurt am Main
Germany