Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 1 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i01.59783

PAPER

Design and Evaluation of Mixed Reality-Based Mobile English Instruction: A Comparative Analysis of Learning Outcomes

Lanlan Fang[1],
Xiujuan Wang[2] (✉),
Liang Zhang[1]

[1]Hunan Mechanical & Electrical Polytechnic, Changsha, China

[2]Hunan Railway Professional Technology College, Zhuzhou, China

p121690@siswa.ukm.edu.my

ABSTRACT

With the deep integration of educational technology and second language acquisition (SLA), mixed reality (MR) technology—combining the immersive qualities of virtual reality (VR) and the contextual relevance of augmented reality (AR)—has provided a novel pathway to address key challenges in traditional mobile English learning, including limited immersion, insufficient interaction, and the disconnection between knowledge and real-world application. This study aims to construct multiple differentiated MR-based mobile English instructional designs and to systematically compare their learning effects across four dimensions: knowledge acquisition, language proficiency, learning engagement, and cognitive load. A quasi-experimental design was employed, in which participants were randomly assigned to one of four groups: a context-immersive MR group, a collaborative inquiry MR group, a personalized adaptive MR group, and a traditional mobile learning control group. An eight-week instructional intervention was implemented. The expected findings indicate that (a) all three MR-based instructional designs perform significantly better than the traditional mobile learning model in vocabulary retention, listening and speaking proficiency, and learning engagement; (b) the context-immersive MR group demonstrates superior performance in contextualized language use and cultural awareness; (c) the collaborative inquiry MR group exhibits distinct advantages in cooperative communication and problem-solving skills; and (d) the personalized adaptive MR group experiences the lowest cognitive load and the highest satisfaction across learners of different proficiency levels. This study contributes to the theoretical development of technology-enhanced SLA, provides empirical support for optimizing MR-based mobile English instructional design, and offers practical guidance for the deeper integration of MR technology into second language education.

KEYWORDS

mixed reality (MR), mobile English instruction, instructional design, learning outcomes, comparative study, second language acquisition (SLA)

Fang, L., Wang, X., Zhang, L. (2026). Design and Evaluation of Mixed Reality-Based Mobile English Instruction: A Comparative Analysis of Learning Outcomes. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(1), pp. 4–18. https://doi.org/10.3991/ijim.v20i01.59783

Article submitted 2025-10-03. Revision uploaded 2025-11-21. Final acceptance 2025-11-25.

1 INTRODUCTION

The deepening of globalization [1, 2] and the widespread adoption of mobile Internet technologies [3, 4] have positioned mobile English learning as one of the core contexts for second language acquisition (SLA). Mobile learning, with its convenience and capacity for fragmented, on-demand access, has effectively transcended the temporal and spatial constraints of traditional classrooms [5]. However, existing models continue to face significant challenges: the lack of immersion makes it difficult for learners to engage with authentic language use scenarios; interaction remains largely limited to one-way human-computer feedback [6]; and a persistent disconnection exists between the transmission of linguistic knowledge and its contextualized application in real-world settings [7]. These limitations collectively hinder the enhancement of learners' integrated language competence.

Mixed reality (MR) technology, through its fusion of virtual and physical environments, integrates the immersive experience of virtual reality (VR) with the contextual relevance of augmented reality (AR), enabling the creation of learning environments that are interactive, perceptible, and immersive [8, 9]. Abstract linguistic knowledge can thereby be transformed into concrete, visualized virtual content, allowing multidirectional interaction among learners, virtual environments, and peers. Such affordances provide essential technological support for addressing the inherent limitations of traditional mobile English instruction.

Although the application of MR technology in the field of education has increased progressively [10], existing research has predominantly focused on isolated application scenarios and lacks systematic instructional design integrating MR technology into mobile English instruction [11]. Moreover, the adaptability and differential effectiveness of diverse instructional designs for varied learning objectives have not been sufficiently validated, making it difficult to establish replicable and generalizable pedagogical paradigms. This limitation constrains the large-scale adoption of MR technology in second language education. The educational use of MR technology has evolved from early-stage feasibility verification toward deeper pedagogical integration across subject domains [9]. In the field of language education, current research has primarily concentrated on three domains. The first concerns vocabulary and grammar instruction, where three-dimensional virtual models and virtual-real overlay annotations are employed to reduce memory difficulty [12]. The second focuses on listening and speaking training, wherein virtual dialogue scenes and character-based role-play environments are constructed to provide immersive oral practice experiences [13]. The third pertains to cross-cultural experiences, in which authentic cultural settings of target-language countries are digitally reconstructed to enhance cultural awareness [14]. Collectively, these studies have verified the positive influence of MR technology in enhancing learning motivation and short-term knowledge acquisition [15]. However, most have explored single-model applications and lack systematic instructional design frameworks.

The implementation of MR-empowered mobile English instruction must be grounded in robust educational theories. Situated learning theory emphasizes the central role of authentic contexts in language acquisition, and the hybrid virtual-physical environments constructed through MR technology provide a tangible embodiment of this theoretical foundation [16]. Constructivist learning theory posits that learners actively construct knowledge through interaction, a process that aligns closely with the multidirectional interactive affordances enabled by MR environments [17]. Furthermore, collaborative learning theory and personalized learning theory offer guidance for differentiated instructional design. The former highlights the role of group interaction in deepening linguistic and cognitive engagement, whereas the latter focuses on addressing individual learner differences and adaptive learning needs [18]. Although existing research has acknowledged these theoretical perspectives, few studies have explored their concrete pedagogical pathways within MR-based mobile instruction environments, resulting in a disjunction between technological application and theoretical grounding.

Drawing on literature analysis and identified research gaps, this study was designed with three primary objectives: first, to construct a series of differentiated MR-based mobile English instructional design frameworks, clarifying their theoretical underpinnings, core components, and implementation processes; second, to conduct a quasi-experimental investigation comparing the learning outcomes of these distinct designs, thereby revealing their respective advantages and suitable application contexts; and third, to examine the moderating effect of individual learner differences on instructional effectiveness, elucidating the underlying mechanisms through which MR technology influences SLA. This study further provides frontline English instructors with three operable MR-based mobile instructional design schemes and practical implementation guidelines to reduce technological barriers, while offering educational technology developers functional optimization recommendations and defining the core functional modules and design directions of MR-supported English instruction platforms.

2 DIFFERENTIATED INSTRUCTIONAL DESIGN FRAMEWORK FOR MR-BASED MOBILE ENGLISH CLASSROOMS

Figure 1 illustrates the proposed research paradigm for MR-based mobile English classroom instructional design. This paradigm is structured around a core "needs-design-evaluation" logic and encompasses three progressive phases: pre-design, design implementation, and post-design. In the pre-design phase, instructional pain points are identified through comprehensive analyses of learner characteristics and investigations into MR-based instructional requirements. In the design implementation phase, three differentiated MR instructional schemes are developed—grounded in theories such as situated learning and collaborative learning—while core strategies such as virtual-real integration and dynamic adaptation are derived. In the post-design phase, iterative optimization and outcome diffusion are achieved through multi-stakeholder feedback and multidimensional evaluation mechanisms. This paradigm provides a systematic framework to support differentiated instructional design.

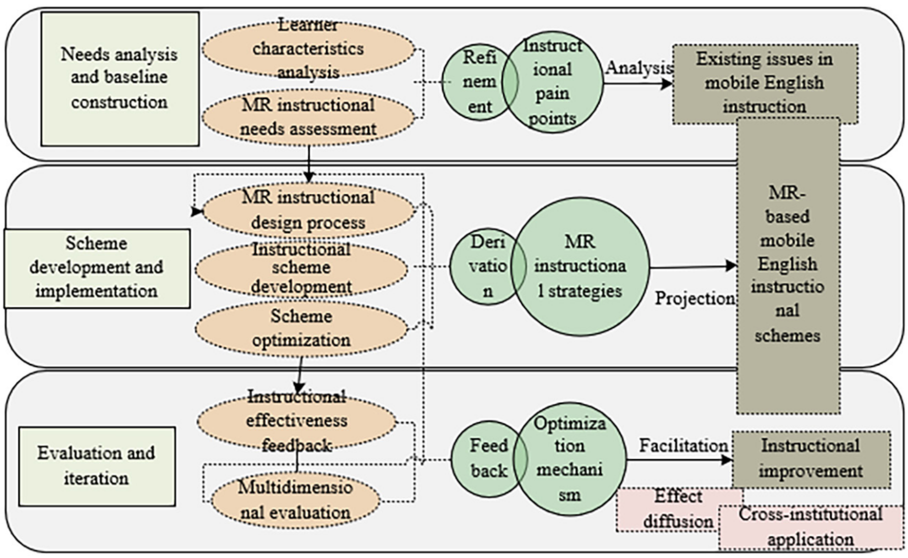

Fig. 1. Research paradigm of MR-based mobile English classroom instructional design

2.1 Scheme I: Context-immersive MR instructional design

The context-immersive MR instructional design is theoretically grounded in situated learning theory. Its core design philosophy lies in utilizing the virtual-real integration capabilities of MR technology to reconstruct authentic language-use scenarios. Through this approach, abstract linguistic knowledge is transformed into perceivable and interactive experiential learning, enabling learners to acquire and produce language naturally through learning-by-doing. This process fosters the development of contextualized language application abilities and intercultural communicative competence. The core elements of this instructional design are achieved through the deep coupling of technological affordances and pedagogical objectives, as shown in Table 1.

The implementation process followed a closed-loop structure of "pre-class preparation-in-class immersion-post-class reflection." In the pre-class phase, preparatory tasks related to the target scenarios were released by the instructor through the MR instructional platform. Learners downloaded scenario data packages via mobile devices and familiarized themselves with core vocabulary, sentence patterns, and relevant cultural contexts using the virtual annotation function. The platform automatically recorded each learner's progress and identified weak points. In the in-class phase, learners entered real-world spaces corresponding to the virtual scenarios. By scanning physical markers with their mobile devices, virtual overlays were triggered. Learners engaged in real-time conversations with virtual characters equipped with natural language interaction capabilities and completed scenario-based tasks. The instructor monitored learners' interaction data through the system dashboard and could intervene via a virtual avatar to provide personalized guidance when necessary. In the post-class phase, learners reviewed recordings of their interaction processes. The platform automatically generated analytical reports on language errors and delivered targeted extended exercises, facilitating the reinforcement and deepening of learning outcomes. This scheme is particularly suitable for English listening and speaking training and the cultivation of intercultural communicative competence, especially for language practice requiring authentic contextual support. The design effectively reduces learners' anxiety in oral expression and enhances their linguistic responsiveness in real-world situations.

Table 1. Core elements of context-immersive MR instructional design

Core Element	Technological Implementation	Instructional Objective
Authentic scene mapping	Real-world settings such as airports, restaurants, and business meeting rooms are digitally modeled using Vuforia-based image recognition technology. Virtual scene overlays are activated when physical markers are scanned through mobile devices.	To construct authentic contexts for language use and reduce the gap between classroom learning and real-world application.
Virtual character interaction	Natural language processing (NLP) technology is employed to enable virtual characters with speech recognition, real-time response, and language correction capabilities, supporting multi-turn conversational interaction.	To provide a low-anxiety communicative environment that strengthens comprehensible input and facilitates the negotiation of meaning.
Scenario-based task sequencing	A hierarchical task chain is designed (basic: information inquiry → intermediate: problem solving → advanced: cross-cultural adaptation), with task difficulty dynamically adjusted according to learner performance.	To progressively enhance learners' language application skills and cultivate strategic awareness for context-sensitive language use.

2.2 Scheme II: Collaborative inquiry MR instructional design

The collaborative inquiry MR instructional design is underpinned by collaborative learning theory and constructivist learning theory. Its core pedagogical concept lies in employing MR technology to construct a shared virtual task environment that transcends physical spatial limitations. Through group role allocation, real-time interaction, and joint problem-solving, learners are guided to deepen linguistic cognition, strengthen cooperative communication, and cultivate critical thinking skills within collective inquiry processes. The core elements of this scheme and their technological-pedagogical alignment are presented in Table 2.

The implementation process followed a three-stage logic of "theme initiation-collaborative inquiry-outcome iteration." During the pre-class phase, inquiry topics were released by the instructor in alignment with instructional objectives. Learners were grouped into teams of four to five members and assigned roles via the MR platform. The system automatically distributed topic-related English literature, case materials, and inquiry task lists. In the in-class phase, teams accessed a shared MR-based virtual task environment through mobile devices. This environment supported real-time synchronous interaction among multiple users. Learners collaborated using tools such as voice communication, virtual annotation, and document sharing to complete tasks, including data synthesis, argument discussion, and solution construction. For instance, in the "cross-cultural business negotiation" theme, each group was required to collaboratively complete tasks such as presenting market research data, designing negotiation scripts, and resolving contentious issues within a virtual negotiation scenario. Throughout the process, the MR platform automatically recorded collaboration data. In the post-class phase, each team produced an English inquiry report or presentation based on the collaborative process and presented it through the platform's virtual display function. Peer evaluation was conducted by other groups through virtual annotations and real-time questioning. The instructor performed a comprehensive evaluation integrating system-recorded data with the quality of the submitted outcomes and provided targeted optimization recommendations. This scheme is particularly suited to English reading extension, collaborative writing, and critical thinking development. By employing group-based inquiry tasks, the design facilitates deep linguistic processing and enhances the logical structuring of English expression. Furthermore, through collaboration and interaction, learners' cross-role communication skills and problem-solving abilities are significantly strengthened.

Table 2. Core elements of collaborative inquiry MR instructional design

Core Element	Technological Implementation	Instructional Objective
Shared virtual tasks	A synchronous virtual environment is constructed based on Unity's multi-user networking technology. The system integrates topic-related virtual scenes, literature resources, and task modules, supporting real-time perspective synchronization among multiple users.	To provide an immersive collaborative context, maintain group goal alignment, and promote focused engagement with the inquiry topic.
Real-time collaboration tools	Integrated functions include real-time voice communication, three-dimensional virtual annotation, and cloud-based document co-editing, supporting English speech-to-text conversion and real-time grammar verification.	To reduce communication barriers, enhance the accuracy of English expression, and improve collaborative efficiency.
Role allocation and dynamic adjustment	A built-in role and permission management module within the MR platform supports dynamic switching of roles during task execution, with automatic recording of task completion data for each participant.	To clarify individual responsibility, encourage complementary strengths, and foster team coordination and adaptive problem-solving skills.

2.3 Scheme III: Personalized adaptive MR instructional design

The personalized adaptive MR instructional design is theoretically grounded in cognitive load theory and personalized learning theory. Its core design philosophy lies in integrating MR technology with learning analytics and dynamic adaptation algorithms to accurately capture individual learner differences and construct personalized learning trajectories within mobile English learning environments. By controlling extraneous cognitive load while maximizing germane cognitive load, efficient and targeted acquisition of linguistic knowledge and skill enhancement are achieved. The core elements of this scheme and their technological-pedagogical coupling logic are presented in Table 3.

Table 3. Core elements of personalized adaptive MR instructional design

Core Element	Technological Implementation	Instructional Objective
Multidimensional learner profiling	The Item Response Theory (IRT) model, Visual-Aural-Read/Write-Kinesthetic (VARK) questionnaire data, and learning analytics are integrated to construct dynamically updated learner profiles from three dimensions: knowledge proficiency, learning style, and skill deficiency.	To precisely identify individual learner differences and provide data-driven support for personalized adaptation.
Dynamic task adjustment	Based on a reinforcement learning algorithm, task difficulty—including parameters such as vocabulary difficulty coefficient and dialogue complexity—task type (e.g., memory-based or application-oriented), and presentation format (e.g., visualized or interactive) are dynamically optimized in conjunction with real-time learning performance data.	To align with learners' zones of proximal development, reduce extraneous cognitive load, and enhance learning efficiency and sense of achievement.
Instant personalized feedback	NLP and speech recognition technologies are integrated to provide real-time pronunciation correction, grammatical error annotation, and individualized problem-solving guidance, accompanied by automatic generation of error attribution analyses.	To promptly address knowledge gaps, standardize linguistic expression, and strengthen learners' autonomous learning capability.

The implementation process was structured as a closed-loop system of "profile construction-dynamic adaptation-feedback iteration." During the pre-class phase, learners completed an English proficiency assessment and a learning style questionnaire through the MR platform. The system integrated these results with historical learning data to generate multidimensional learner profiles, identifying individual weaknesses and learning preferences. Based on the profiles, personalized learning schemes were automatically generated, specifying adaptive task types, difficulty gradients, and learning paces. In the in-class phase, learners accessed the personalized MR learning module, where real-time behavioral data were tracked to adjust the instructional content and delivery format dynamically. For instance, learners with weak vocabulary mastery were assigned MR-based visualized memory tasks; those with inaccurate pronunciation were guided through one-on-one virtual pronunciation correction sessions with AI tutors; and those with grammatical confusion were provided with integrated virtual-real sentence deconstruction demonstrations. Meanwhile, extraneous cognitive load was controlled by simplifying redundant virtual elements and sequencing complex information step by step, ensuring that each learning task remained within the learner's zone of proximal development. In the post-class phase, the MR platform automatically generated detailed learning reports that quantified learners' improvement in weak areas, mastery levels, and overall skill development trajectories. Based on these reports, the system recommended adaptive extension resources. Learners could autonomously select review modules, while the platform supported learning resumption from prior progress points and provided targeted reinforcement training. This scheme is particularly applicable to English vocabulary and grammar reinforcement as well as personalized listening and speaking practice. It is especially effective in classroom contexts where learners exhibit heterogeneous proficiency levels, addressing the "one-size-fits-all" limitation of traditional mobile learning and enabling precise and efficient development of foundational language skills.

3 EXPERIMENTAL DESIGN AND RESEARCH METHODOLOGY

A quasi-experimental design was employed to ensure the validity and comparability of the experimental results through the explicit definition of independent, dependent, and controlled variables. The independent variable is the mobile English classroom instructional design scheme, which includes four levels: (a) context-immersive MR instructional design; (b) collaborative inquiry MR instructional design; (c) personalized adaptive MR instructional design; and (d) the traditional mobile learning model. The core distinctions among these four conditions lie in their respective modes of technological integration, forms of instructional interaction, and logic of learning task design. The dependent variables focus on the multidimensional evaluation of learning outcomes across four principal dimensions: (a) knowledge acquisition, quantified through vocabulary and grammar tests; (b) language proficiency, assessed through a combination of component-specific and integrated evaluations, including listening, speaking, reading, and writing tests; (c) learning process, measured through indicators of learning engagement and cognitive load; and (d) learning attitudes, encompassing technology acceptance, learning satisfaction, and learning motivation.

As for the controlled variables, the instructional content was standardized across all experimental groups, with three units selected from the College English textbook for non-English majors at a comprehensive university to ensure material consistency. The instructional duration was fixed at eight weeks, with two class sessions per week, resulting in an equal total teaching time across groups. The same instructor—who possessed over ten years of university English teaching experience and had received professional training in MR-based pedagogy—was assigned to all groups to minimize the influence of teaching style variability. In addition, uniform learning devices were provided, consisting of mobile terminals equipped with Android 12.0 or higher systems supporting ARCore/ARKit, pre-installed with the customized MR instructional platform and the traditional mobile learning application. This ensured that no discrepancies in device performance affected the experimental outcomes.

Figure 2 illustrates the interactive logic between the instructor and learners across the three instructional stages—pre-class, in-class, and post-class—with the MR-based mobile instructional platform serving as the central medium. In the pre-class stage, the instructor distributed MR-based preparatory resources and analyzed diagnostic data, while learners engaged in virtual preview activities. During the in-class stage, the instructor adjusted instructional priorities based on pre-class data, organized MR scenario-based interactions, and provided real-time feedback, whereas learners participated in immersive communicative training. In the post-class stage, personalized MR review tasks were released, enabling learners to reflect on their interactive experiences and submit extended learning outputs. The figure explicitly delineates the core MR technological functions and data flow pathways across all stages, providing a procedural reference for the implementation of experimental interventions.

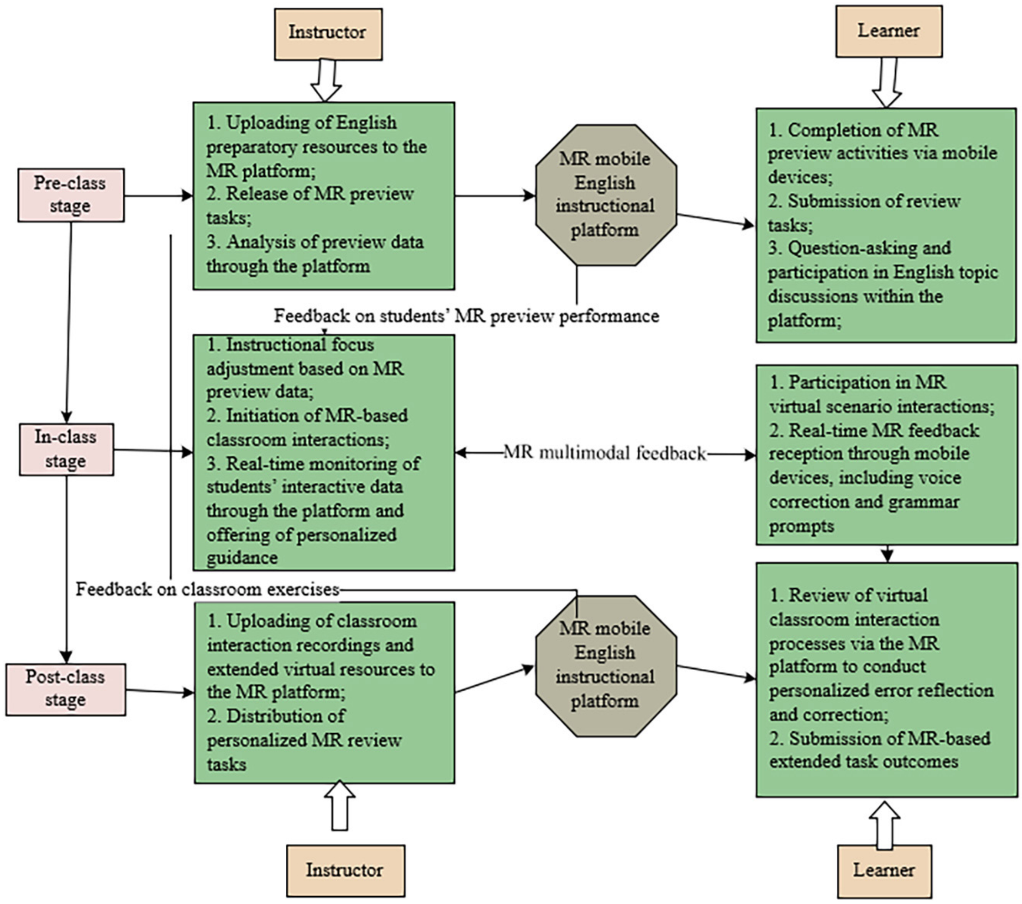

Fig. 2. Implementation process of mobile English classroom instruction based on MR technology

4 EXPECTED RESULTS AND ANALYSIS

To examine the differential effects of the three MR-based instructional design schemes on the acquisition of core English knowledge, inter-group comparisons were conducted focusing on two fundamental dimensions—vocabulary and grammar. The simplified results are presented in Table 4. As shown in the table, after an eight-week intervention, the mean posttest scores of all three MR experimental groups in vocabulary, grammar, and overall knowledge acquisition were significantly higher than those of the control group. Specifically, the vocabulary dimension showed an average improvement of 22.8%–30.5%, while grammar improved by 18.2%–24.9%. In contrast, the control group exhibited much smaller gains of 11.7% in vocabulary and 9.2% in grammar. These findings indicate that MR technology, through its visualized and interactive presentation of instructional content, substantially reduced the cognitive difficulty of language knowledge acquisition, outperforming the unidirectional transmission characteristic of traditional mobile learning models. Among the experimental groups, the personalized adaptive MR group achieved the highest overall performance, with statistically significant differences compared to the other groups ($p < 0.05$). This superiority can be attributed to the scheme's precise adaptation to learners' individual weaknesses and its targeted reinforcement mechanisms. In contrast, the context-immersive MR group and the collaborative inquiry MR group did not differ significantly ($p > 0.05$), likely due to the fact that their core designs emphasized scenario-based interaction and collaborative exploration rather than targeted knowledge consolidation. In summary, the results demonstrate that MR-based instructional design exerts a significant positive effect on English knowledge acquisition. Among the three MR schemes, the personalized adaptive MR design exhibited the strongest enabling effect.

Table 4. Effects of different instructional design schemes on English knowledge acquisition (posttest scores, M ± SD, n = [X])

Dimension	Context-Immersive MR Group	Collaborative Inquiry MR Group	Personalized Adaptive MR Group	Control Group	F-Value	p-Value	Post hoc Comparison (Bonferroni)
Vocabulary acquisition (vocabulary size + application)	76.2 ± 6.4	75.5 ± 6.6	81.9 ± 6.1	64.8 ± 7.0	41.39	<0.001	Personalized group > context group = collaborative group > control group
Core grammar application	74.4 ± 6.3	73.9 ± 6.5	78.4 ± 5.9	62.1 ± 7.1	35.76	<0.001	
Total knowledge acquisition score	75.3 ± 6.2	74.7 ± 6.4	79.6 ± 5.8	62.5 ± 6.9	44.82	<0.001	

To highlight the core distinctions among instructional schemes, the language proficiency evaluation in this study was simplified to include four primary components—listening, speaking, reading, and writing—along with an overall composite score. Additionally, interaction effects between learners' proficiency levels were examined. The results are presented in Table 5. As shown in the table, the total language proficiency scores for all MR experimental groups (ranging from 74.8 to 78.3) were significantly higher than those of the control group (60.2), indicating that the performance gains of MR-assisted instruction were substantially greater than those of traditional mobile learning. These findings confirm the facilitating role of MR technology in enhancing overall language competence. As shown in the detailed comparison, the context-immersive MR group achieved the highest mean scores in listening ($M = 78.1$) and speaking ($M = 78.7$) with statistical significance ($p < 0.05$). This outcome aligns with the design emphasis of this scheme on authentic communicative environments, which effectively strengthened contextualized language input and output.

In addition, the collaborative inquiry MR group demonstrated superior performance in reading ($M = 77.2$) and writing ($M = 77.8$) ($p < 0.05$), a result attributable to the intensive training in linguistic logic and expression accuracy fostered through collaborative processes. The personalized adaptive MR group exhibited balanced improvement across all skill areas, with no single domain achieving the highest score, yet exhibiting stable overall improvement. The interaction effect analysis revealed that low-proficiency learners benefited most from the personalized adaptive group, whereas high-proficiency learners gained more from the collaborative inquiry group. These findings suggest that instructional design should be calibrated to learners' initial proficiency levels. In summary, MR-based instructional design was found to promote differentiated enhancement across language dimensions. The context-immersive scheme was most effective for listening and speaking development, and the collaborative inquiry scheme was best suited for reading and writing advancement. In addition, matching instructional design to learners' proficiency levels is critical for optimizing learning outcomes in MR-enhanced mobile English instruction.

Table 5. Effects of different instructional design schemes on English language proficiency (posttest scores, M ± SD, n = [X])

Dimension	Context-Immersive MR Group	Collaborative Inquiry MR Group	Personalized Adaptive MR Group	Control Group	F-Value	p-Value	Post hoc Comparison	Interaction Effect (Proficiency × Scheme)
Listening ability	78.1 ± 6.0	73.9 ± 6.4	74.5 ± 6.2	64.9 ± 6.8	39.57	<0.001	Context group > personalized group = collaborative group > control group	p = 0.024 < 0.05
Speaking ability	78.7 ± 5.7	73.4 ± 6.5	74.1 ± 6.0	62.3 ± 7.1	44.93	<0.001		p = 0.017 < 0.05
Reading ability	73.8 ± 6.2	77.2 ± 5.8	74.3 ± 6.1	64.2 ± 6.9	37.82	<0.001	Collaborative group > personalized group = context group > control group	p = 0.022 < 0.05
Writing ability	73.6 ± 6.3	77.8 ± 5.7	73.9 ± 6.2	62.5 ± 7.0	41.68	<0.001		p = 0.018 < 0.05
Total language proficiency score	76.0 ± 5.9	75.6 ± 6.0	74.8 ± 6.1	60.2 ± 6.8	47.95	<0.001	Context group = collaborative group > personalized group > control group	p = 0.026 < 0.05
Improvement rate of low-proficiency learners	25.7%	22.3%	28.3%	7.9%	——	——	Personalized group > context group > collaborative group > control group	——
Improvement rate of high-proficiency learners	20.3%	27.1%	21.5%	8.1%	——	——	Collaborative group > context group > personalized group > control group	——

To examine the differentiated performance of key indicators related to the learning process, the evaluation dimensions were streamlined to include learning engagement and cognitive load. The results are presented in Table 6. As shown in the table, in terms of learning engagement, both the context-immersive MR group (M = 77.3) and the collaborative inquiry MR group (M = 76.8) achieved significantly higher total scores than the other groups ($p < 0.05$), with no statistically significant difference between them. A further breakdown revealed that the context-immersive MR group achieved the highest level of emotional engagement (M = 78.0), while the collaborative inquiry MR group led in behavioral engagement (M = 77.1). These findings validate the effectiveness of virtual interaction and collaborative tasks in enhancing learner participation. In terms of cognitive load, the personalized adaptive MR group demonstrated the most favorable outcomes, exhibiting the lowest subjective cognitive load (M = 3.2) and the shortest objective task duration (M = 38.5 minutes) ($p < 0.05$). Although the context-immersive MR group showed a slightly higher cognitive load, its values remained significantly below those of the control group. This result suggests that the amount of information processed in realistic scenarios did not exceed learners' cognitive capacity. In summary, the findings indicate that the context-immersive and collaborative inquiry MR designs primarily enhance learning engagement, while the personalized adaptive MR design excels in cognitive load regulation, achieving the intended differentiated alignment of process-oriented objectives.

To systematically verify the multidimensional enhancement effects of different MR instructional design schemes on English learning, an inter-group comparative analysis was conducted across four key dimensions: vocabulary mastery, listening comprehension, oral fluency, and learning engagement. As shown in Figure 3, in the vocabulary mastery dimension, the personalized adaptive MR group achieved the highest performance, with 65% of learners scoring within the 80–89 range, a proportion significantly higher than that of the other groups ($p < 0.05$). This result demonstrates the scheme's core advantage in precisely targeting individual lexical weaknesses. In the listening comprehension and oral fluency dimensions, the context-immersive MR group exhibited the strongest performance, with 60% of learners scoring within the 80–89 range for listening and 65% for speaking. These outcomes can be attributed to the scheme's immersive communicative environments, which strengthened contextualized input and output. Regarding learning engagement, the collaborative inquiry MR group achieved a 60% distribution within the 80–89 range because the collaborative inquiry tasks encouraged deeper engagement. In contrast, the traditional mobile learning group exhibited substantially higher proportions of learners in the lower 60–79 score range across all dimensions, highlighting its deficiencies in knowledge acquisition, skill development, and engagement. In summary, MR-based instructional designs demonstrated significant positive effects across all dimensions of English learning, with distinct advantages aligned to specific learning outcomes. The personalized adaptive MR scheme was found to be most effective for precise vocabulary acquisition, the context-immersive MR scheme excelled in listening and speaking training within contextualized environments, and the collaborative inquiry MR scheme was best suited for deep interactive learning. These patterns provide empirical evidence for the targeted application of MR technology in achieving diverse pedagogical objectives in mobile English instruction.

Table 6. Effects of different instructional design schemes on learning process outcomes (M ± SD, n = [X])

Dimension		Context-Immersive MR Group	Collaborative Inquiry MR Group	Personalized Adaptive MR Group	Control Group	F-Value	p-Value	Post hoc Comparison (Bonferroni)
Learning engagement	Behavioral engagement (interaction frequency)	76.6 ± 6.1	77.1 ± 6.0	72.3 ± 6.2	61.3 ± 7.0	42.15	<0.001	Context group = collaborative group > personalized group > control group
	Emotional engagement (enjoyment)	78.0 ± 5.8	75.6 ± 6.1	72.4 ± 6.1	62.8 ± 6.9	46.93	<0.001	
	Total learning engagement	77.3 ± 5.9	76.8 ± 6.0	72.2 ± 6.1	61.6 ± 6.9	48.75	<0.001	
Cognitive load	Subjective cognitive load (Paas scale)	4.1 ± 0.8	4.5 ± 0.9	3.2 ± 0.7	5.8 ± 1.0	40.79	<0.001	Personalized group < context group < collaborative group < control group
	Objective task duration (min)	42.3 ± 4.5	45.1 ± 4.8	38.5 ± 4.2	52.6 ± 5.3	37.95	<0.001	

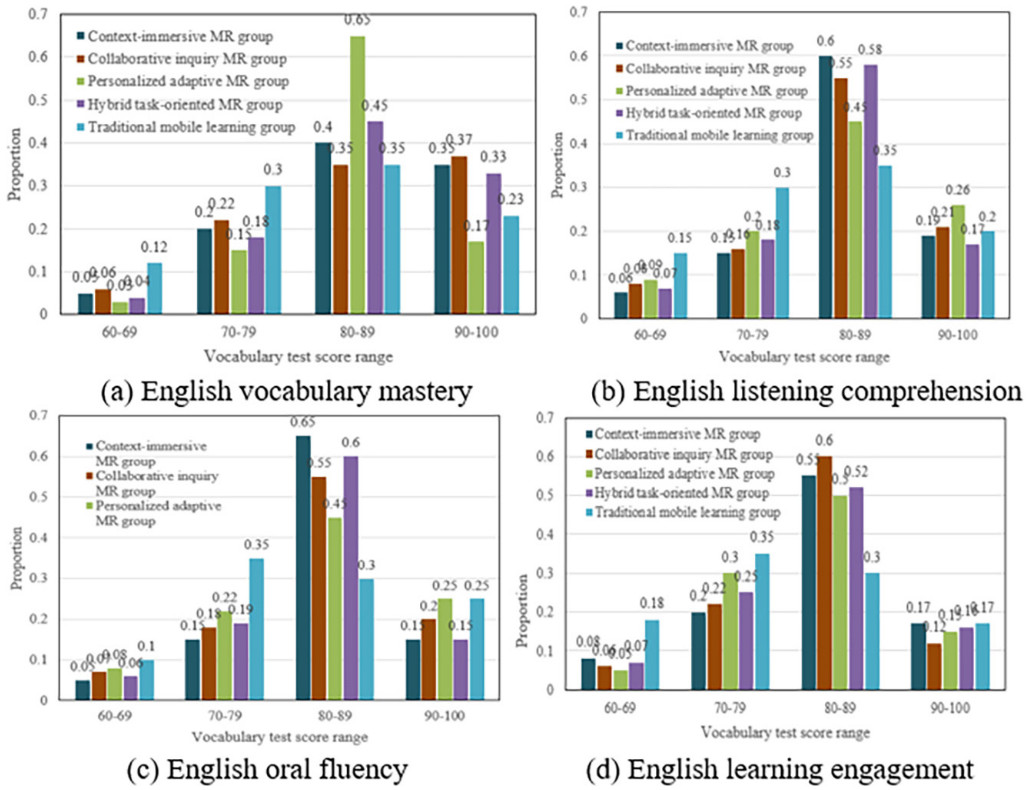

Fig. 3. Comparative distribution of multidimensional learning outcomes between differentiated instructional designs and the traditional model

5 CONCLUSION

This study was conducted to explore the deep integration of MR technology and English instruction through the development of four differentiated instructional design schemes: the context-immersive, collaborative inquiry, personalized adaptive, and hybrid task-oriented MR teaching designs. Using a quasi-experimental framework, the effectiveness of these MR-based instructional designs was systematically verified across four core dimensions—English core knowledge acquisition, comprehensive language ability, learning process, and learning attitude. The key findings revealed that all MR experimental groups performed significantly better than the traditional mobile learning group across all evaluated dimensions, demonstrating distinct dimension-specific advantages. The personalized adaptive MR scheme, through precise adaptation to individual learners' weaknesses, achieved the best results in vocabulary and grammar acquisition. The context-immersive MR scheme, by constructing authentic communicative environments, proved to be the most effective for enhancing listening comprehension and oral fluency. The collaborative inquiry MR scheme, grounded in group-based exploratory tasks, effectively strengthened reading and writing abilities while improving learning engagement. The hybrid task-oriented MR scheme achieved a balanced enhancement across multiple dimensions. The theoretical value of this research lies in its extension of the practical connotations of SLA theory and situated learning theory within MR-based instructional environments. A logical framework of "technological empowerment—dimensional adaptation—outcome optimization" was constructed for the MR-based

English instruction. From a practical perspective, four implementable MR instructional design schemes and their corresponding evaluation tools were developed, providing precise methodological guidance for aligning MR applications with specific pedagogical objectives. This approach effectively addresses the core limitations of traditional mobile English instruction, namely, insufficient immersion, weak interactivity, and lack of personalization.

Several limitations were identified in this study that warrant further refinement in subsequent research. First, the experimental sample did not encompass a sufficiently diverse range of higher education institutions or learners with varying levels of English proficiency, which may constrain the generalizability of the findings. Second, the duration of the intervention was insufficient to fully assess the long-term impact and transferability of knowledge and skills derived from the MR-based instructional designs. Additionally, the influence of different hardware configurations on the MR learning experience was not examined. Third, the study primarily focused on the overall effectiveness of each instructional scheme, while the in-depth exploration of learner-level differences and the underlying interactive mechanisms within MR environments remained limited. Future research could be advanced in three directions. First, the sample scope should be expanded to include institutions of different tiers and learner populations with varying English proficiency levels. Multicenter experimental designs should be employed to enhance the external validity of the conclusions. Second, the intervention period should be extended to an entire academic semester to enable longitudinal tracking of sustained learning outcomes. Concurrently, MR technology should be optimized to ensure compatibility with mid- and low-end mobile devices, thereby reducing barriers to adoption. Third, integration with artificial intelligence technologies should be deepened to refine the logic of personalized adaptation. Dynamic monitoring indicators of learning engagement and cognitive load should be further elaborated, while additional moderating variables may be introduced to examine the optimization pathways of MR-based instruction under multifactor interaction, thereby advancing the scaled and precision-oriented application of MR technology in the domain of English language education.

6 ACKNOWLEDGMENTS

This paper is the research outcome of the 2022 Hunan Provincial Vocational Education Teaching Reform Research Project titled Construction and Practical Research on the Digital Learning Ecosystem of English in Higher Vocational Colleges (Grant No.: ZJGB2022333), led by Principal Investigator Lanlan Fang.

7 REFERENCES

[1] W. Qian, "Enhancing university English teaching through mobile interactive applications from an educational technology perspective," *International Journal of Interactive Mobile Technologies*, vol. 18, no. 20, pp. 51–65, 2024. https://doi.org/10.3991/ijim.v18i20.51885

[2] F. Khodabandeh, A. Bagheri, and M. I. Mobarakeh, "Augmented reality's role in EFL learning: Enhancing language skills in clothing and style through flipped and blended methods," *Virtual Reality*, vol. 29, no. 2, pp. 1–16, 2025. https://doi.org/10.1007/s10055-025-01151-y

[3] H. Chang and J. Zhao, "Enhancing translation teaching efficiency and learning experience through mobile technology applications," *International Journal of Interactive Mobile Technologies*, vol. 19, no. 6, pp. 75–89, 2025. https://doi.org/10.3991/ijim.v19i06.54703

[4] S. Lu, "Interactive virtual reality environments for English language learning: Immersive technologies to enhance language acquisition and communication skills," *Journal of Computational Methods in Sciences and Engineering*, vol. 25, no. 4, pp. 3093–3106, 2025. https://doi.org/10.1177/14727978251318803

[5] R. Kaplan-Rakowski and T. Thrasher, "The impact of high-immersion virtual reality and interactivity on vocabulary learning," *British Journal of Educational Technology*, vol. 56, no. 6, pp. 2647–2670, 2025. https://doi.org/10.1111/bjet.13603

[6] G. Andreou and M. Gkantaki, "An eye-tracking study on text comprehension while listening to music: Preliminary results," *Applied Sciences (Switzerland)*, vol. 15, no. 7, pp. 1–12, 2025. https://doi.org/10.3390/app15073939

[7] Y. Xia, S.-Y. Shin, and K.-S. Shin, "Designing personalized learning paths for foreign language acquisition using big data: Theoretical and empirical analysis," *Applied Sciences (Switzerland)*, vol. 14, no. 20, pp. 1–20, 2024. https://doi.org/10.3390/app14209506

[8] R. Wafiqoh, F. Apriani, and I. Zaliman, "The effectiveness of virtual reality in enhancing students' reflective abstraction in the independent curriculum," *Mathematical Modelling of Engineering Problems*, vol. 12, no. 8, pp. 2855–2862, 2025. https://doi.org/10.18280/mmep.120826

[9] O. Shkvyr, H. Dudchak, N. Kazakova, O. Polianovska, and N. Sivak, "Mathematical modeling of information technology integration in digital education: A regional perspective," *Ingénierie des Systèmes d'Information*, vol. 28, no. 3, pp. 603–610, 2023. https://doi.org/10.18280/isi.280308

[10] E. Christou, A. Parmaxi, and M. Christoforou, "Implementation and application of extended reality in foreign language education for specific purposes: A systematic literature review," *Universal Access in the Information Society*, vol. 24, no. 3, pp. 2061–2076, 2025. https://doi.org/10.1007/s10209-025-01191-w

[11] K.-W. Su, H.-J. Chen, and C.-T. Chen, "Development and evaluation of a mixed reality system for health education on neck stiffness and soreness," *International Journal of Industrial Ergonomics*, vol. 109, pp. 1–12, 2025. https://doi.org/10.1016/j.ergon.2025.103779

[12] M. Alfadil, "Immersive virtual reality: A novel approach to second language vocabulary acquisition in K-12 education," *Sensors*, vol. 24, no. 22, pp. 1–15, 2024. https://doi.org/10.3390/s24227185

[13] J. Wang, D. Guo, and H. Liu, "Where to learn: Embodied perception learning planned by vision-language models," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 17, no. 5, pp. 1050–1060, 2025. https://doi.org/10.1109/TCDS.2025.3539665

[14] D. Monteiro, T. Ma, Y. Li, Z. Pan, and H.-N. Liang, "Cross-cultural factors influencing the adoption of virtual reality for practical learning," *Universal Access in the Information Society*, vol. 23, no. 3, pp. 1203–1216, 2024. https://doi.org/10.1007/s10209-022-00947-y

[15] M. Vallarino, R. Cersosimo, I. Torre, and G. Vercelli, "Effects of using depth and motion in virtual reality to enhance language learning," *Virtual Reality*, vol. 29, no. 3, pp. 1–14, 2025. https://doi.org/10.1007/s10055-025-01203-3

[16] L. Gabora, "The cultural evolution of socially situated cognition," *Cognitive Systems Research*, vol. 9, nos. 1–2, pp. 104–114, 2008. https://doi.org/10.1016/j.cogsys.2007.05.004

[17] H. Robles, M. Jimeno, K. Villalba, I. Mardini, C. Viloria-Nuñez, and W. Florian, "Design of a micro-learning framework and mobile application using design-based research," *PeerJ Computer Science*, vol. 9, pp. 1–31, 2023. https://doi.org/10.7717/peerj-cs.1223

[18] H. Quené, R. Orr, and D. Van Leeuwen, "Phonetic similarity of /s/ in native and second language: Individual differences in learning curves," *Journal of the Acoustical Society of America*, vol. 142, no. 6, pp. EL519–EL524, 2017. https://doi.org/10.1121/1.5013149

8 AUTHORS

Lanlan Fang Associate Professor. She graduated from Xiangtan University in 2004, majoring in English. She received a master's degree from Hunan Normal University in 2015. She is currently an English teacher at the Hunan Mechanical & Electrical Polytechnic. Her research interests include English teaching, information-based teaching, and vocational education internationalization (E-mail: fanglanlan1982@163.com).

Xiujuan Wang is an Associate Professor. She received the Master's degree from Shanghai Maritime University, P.R. China. She is currently a Lecturer in Hunan Railway Professional Technology College, China. Her research interests include English teaching, information-based teaching and second language acquisition, and international studies (E-mail: p121690@siswa.ukm.edu.my).

Liang Zhang is an Associate Professor. She received the Master's degree from Central South University, P.R. China. Her research interests include vocational English teaching, and vocational education internationalization (E-mail: lenazhang103@126.com).