WISVORA

International Theory and Practice in Humanities and Social Sciences

2026 Volume3,Issue4 ISSN 3078-4387

The Construction of Metaverse English Classroom: The Multi-Layer Immersive Language Learning Environment Framework

Manman Li[1], Tiantian Xie[2*]

[1] *Department of Foreign Languages, Taiyuan Normal University, Jinzhong, 030619,Shanxi, China.*

[2] *Faculty of Education, University of Malaya, Kuala Lumpur 50603, Malaysia*

Accepted

2026-08-12

Keywords

MILLE; Metaverse; Metaverse English Classroom

Corresponding Author

Tiantian Xie

https://doi.org/10.70693/itphss.v3i4.669

Abstract

With the rapid advancement of digital technologies, the metaverse has introduced new possibilities for English education. However, existing studies largely focus on isolated technological applications, lacking a systematic framework that integrates pedagogical and technological elements. To address this gap, this study proposes the MILLE (Metaverse-based Immersive Language Learning Environment) framework, a multi-layer model designed to support immersive, interactive, and adaptive English learning. Grounded in constructivism, situational cognition, and immersive learning theory, the framework integrates key components such as interaction, collaboration, assessment, and support mechanisms. To provide preliminary empirical support, a questionnaire-based study was conducted with 120 participants, including both students and teachers. The results demonstrate satisfactory reliability and validity of the measurement model and reveal generally positive user perceptions in terms of usefulness, immersion, and interaction. Overall, the findings suggest that the MILLE framework is both theoretically grounded and practically feasible. This study contributes by bridging conceptual framework development with empirical validation and offers a structured approach for designing metaverse-based English learning environments.

1. Introduction

The ongoing advancement of globalization has placed increasing demands on individual' comprehensive English proficiency. However, traditional English teaching continues to face significant limitations. Instructional practices are often dominated by grammar-focused and examination-oriented approaches, resulting in three major challenges: limited opportunities for authentic communication, insufficient cultural immersion, and a lack of personalized feedback(Boumová, 2008). Consequently, learners tend to perceive English as a static academic subject rather than a dynamic communicative tool, which constrains both motivation and practical language use(Sharma & Joshi, 2024).

Although computer-assisted language learning (CALL) and mobile-assisted language learning (MALL) have been widely adopted to address these issues, they remain largely confined to two-dimensional interaction environments(Nguyen & Nguyen, 2024). In contrast, the metaverse, as an emerging paradigm of the next-generation internet, integrates technologies such as extended reality and virtual interaction, enabling immersive, real-time, and socially interactive learning experiences(Efgivia et al., 2021). This shift offers new opportunities to enhance presence, engagement, and contextual authenticity in language learning(Hwang & Lee, 2024).

Despite these advancements, existing research has predominantly focused on isolated technological features rather than on the systematic integration of pedagogical principles and technological affordances(Kirshner & Whitson, 2021). This has led to fragmented applications and a lack of comprehensive frameworks to guide the design and implementation of metaverse-based English learning environments. In particular, there is a growing need for models that integrate immersion, interaction, collaboration, and assessment into a coherent system(Lui & Goel, 2022).

To address this gap, this study proposes the MILLE framework, which aims to provide a structured and theoretically grounded model for metaverse-based English learning. The framework is informed by established educational theories and integrates multiple functional layers to support a comprehensive learning ecosystem.

In addition to its theoretical contribution, this study incorporates an empirical component to provide preliminary validation of the proposed framework. Specifically, questionnaire data are used to assess the reliability and validity of the measurement model and to examine users' perceptions of the framework.

Accordingly, this study addresses the following research questions:

RQ1: What are the key components and structure of the proposed MILLE framework?

RQ2: Does the empirical data provide support for the reliability and validity of the framework?

RQ3: How do users perceive the MILLE framework in terms of usefulness, immersion, and interaction?

2. Theoretical Basis

This chapter mainly explains the theoretical basis of the study, and its theoretical framework diagram is shown in Figure 1.

The application of immersive virtual simulation technology in English education is not a simple embedding of cutting-edge technology, but a comprehensive presentation and in-depth practice of a variety of modern educational theories in the digital learning space. First, constructivism emphasizes learners' active participation, interactive collaboration, and knowledge construction in meaningful situations(Rossetti, 2023). The real interactive space built by the metaverse is highly consistent with this concept. Through shared scenarios such as virtual airport inquiries and business negotiations, learners can explore pragmatic strategies in teamwork and problem-solving, thereby effectively transforming language knowledge into practical communication skills. Second, situational cognition theory states that the acquisition of knowledge and skills must be embedded in real or quasi-real situations, internalized and transferred through "use in contexts"(Guo, 2022). With its high degree of realism and operability, the metaverse classroom provides an ideal technical field for this theory. Whether it's a virtual museum tour or a cross-cultural conference simulation, the metaverse places language learning in dynamic social practices, greatly compensating for the shortcomings of traditional classroom situations. In addition, based on Kolb's experiential learning theory, the metaverse constructs a complete closed loop of concrete experience, reflective observation, abstract conceptualization, and active experimental learning. It not only provides an immersive sensory experience, but also uses AI instant feedback and intelligent data analysis to assist reflection, promote concept induction through teaching assistant guidance, and then validate strategies in new tasks, thereby greatly improving learning efficiency.

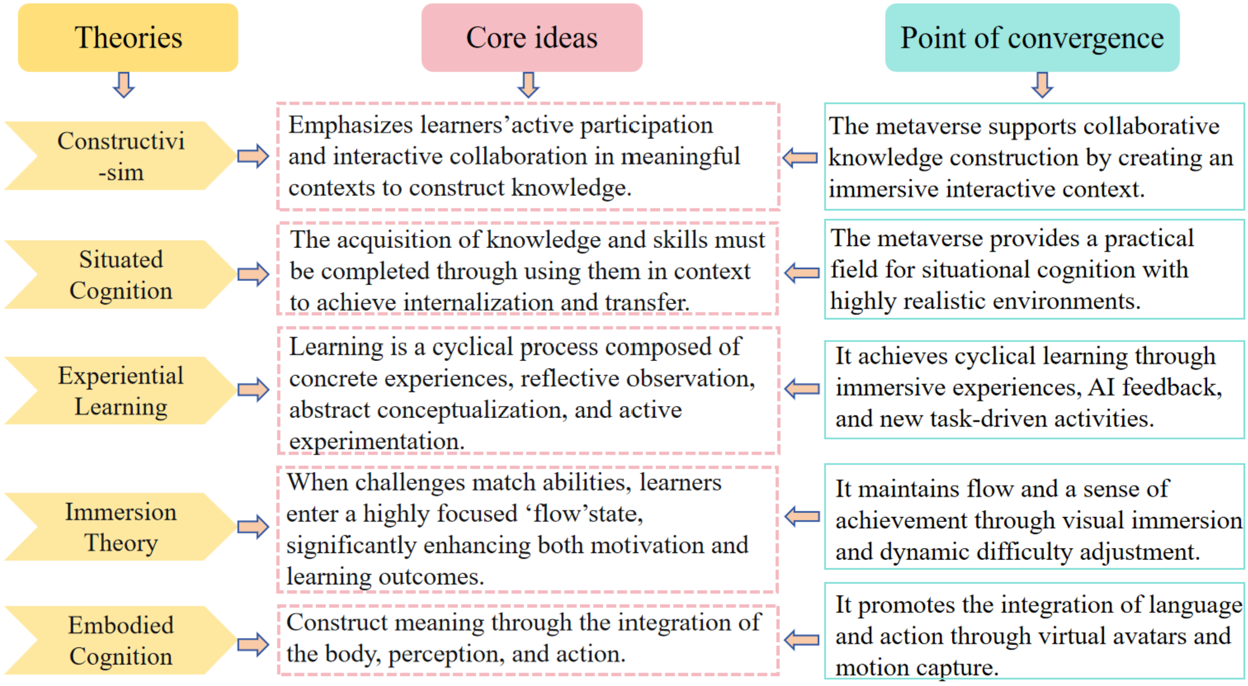

Figure 2.1: Theoretical Framework Diagram

From a psychological perspective, Çiksenmiha's Flow Theory further reveals the advantages of the metaverse. When challenges are matched with abilities, learners are prone to a state of highly focused "flow"(Chen, 2024). Through visual immersion, real-time interaction, and task grading mechanisms, the metaverse can effectively maintain learners' continuous engagement, give them a sense of accomplishment in pragmatic output, and significantly enhance learning motivation. Finally, embodied cognition theory provides a new perspective for understanding the educational value of the metaverse. This theory argues that cognition arises from physical experience and is an integration of body, perception, and action(Therrien, 2023). Compared with the static input of traditional classrooms, the metaverse provides a breakthrough path for embodied participation with the help of virtual avatars, motion capture and spatial interaction technologies(Frost et al., 2022). Learners can not only use language, but also interact with the virtual environment through gestures, gestures and cultural etiquette (e.g., waving in a virtual restaurant to order food)(Farina, 2021). This practice of body-body integration greatly promotes deep language processing and cross-cultural semantic construction(Yagcioglu, 2023). To sum up, the metaverse English classroom breaks through the temporal and spatial limitations of traditional teaching, realizes the organic synergy of constructivism, situational cognition, experiential learning, immersion theory and embodied cognition at the theoretical level, and builds a highly immersive, embodied and interactive comprehensive learning ecology.

3. Related Work

3.1 Limitations of Current Language Learning Platforms

Online English learning, with its flexibility and abundance of resources, breaks through the time and space limitations of traditional education and profoundly shapes modern learning methods. However, while embracing these advantages, we must carefully examine the multidimensional challenges it faces in terms of teaching effectiveness and accessibility, which are rooted in various aspects such as environment, psychology, technology, and pedagogy(Wu, 2025).

Firstly, in terms of environment and interaction, online teaching has noticeable shortcomings. Distractions in the home environment make it difficult for learners to maintain focus, and for younger students, the lack of classroom discipline further increases the difficulty of learning management(Choi & Chang, 2025). More importantly, the absence of face-to-face communication in an online setting can easily lead to feelings of isolation, weakening students' sense of belonging to the learning community and subsequently reducing their willingness to engage in collaborative learning(Parveen et al., 2025). This weak interaction mechanism not only affects the emotional connection between teachers and students and among peers but also limits teachers' ability to promptly grasp students' learning progress and behavior, making it even harder to identify and address negative emotions such as anxiety or frustration in time.

The shortcomings in environment and interaction further have a chain effect on students' psychological state and learning motivation(Chiu et al., 2021). Common feelings of frustration, nervousness, and fear when starting language learning are harder to alleviate effectively in online classes without teacher guidance in person, easily transforming into a negative attitude toward English learning(Peng & Fu, 2021). Many students show a noticeable decline in motivation when transitioning from offline to online learning; this is not primarily due to relaxed teacher supervision but rather due to difficulty adapting to the new teaching format. Since most students have limitations in self-regulation, only a few can maintain proactive learning without external support, while most learners still rely on continuous attention and interactive encouragement from teachers to sustain their enthusiasm for learning(Bahari, 2023).

Moreover, the effectiveness of online English learning is significantly constrained by individual differences and external resource conditions(Van et al., 2021). Differences in students' self-learning ability, self-discipline, and personality traits (such as confidence) directly impact their performance in an online environment. On the other hand, online learning heavily depends on stable technological devices and a reliable internet connection, which are not equally accessible to all families. Issues such as a lack of electronic devices, poor internet connectivity, and insufficient digital literacy at home not only hinder the smooth progress of learning but also invisibly exacerbate concerns regarding educational equity(Purwantoro et al., 2021). Additionally, for language learning that heavily relies on interaction and practice, the online environment struggles to create the 'immersive' atmosphere of a physical classroom that encourages free expression and tolerates mistakes, which in turn limits the effective improvement of language skills from a pedagogical standpoint(Ariffin et al., 2021).

3.2 Immersive Language Learning and Metaverse English Classrooms

With the rapid popularization of immersive technologies represented by VR/AR in the field of education, language learning is transitioning from traditional classrooms to immersive, interactive, and customizable intelligent learning spaces. Immersive virtual simulation technology relies on the integrated application of high-performance computing, human-computer interaction, sensor control, and artificial intelligence to create realistic environments through simulation and 3D presentation(Hein et al., 2021). This allows learners to experience visual, auditory, and even tactile modalities using headsets, gesture recognition devices, and haptic feedback systems, thereby overcoming the limitations of traditional language classrooms in scenario creation, communicative authenticity, and cultural representation, achieving a language learning experience that feels like being in a real context(Weng et al., 2024). Research shows that such immersive learning environments can not only significantly enhance learners' engagement and motivation but also activate various learning styles through experiential activities, providing learners with

opportunities for active exploration and interactive expression(Dhimolea et al., 2022).
The core value of immersive learning lies in its three main characteristics: immersiveness, interactivity, and imagination(Fernandes et al., 2023). Immersiveness allows learners to move beyond merely observing language situations and establish direct perceptual connections with virtual spaces through sensing devices, performing real-time pragmatic expressions in simulated tasks such as airport inquiries, restaurant ordering, and international meetings(Beck et al., 2023). Interactivity enables learners to actively manipulate virtual objects, converse with virtual characters in real time, and choose communication strategies, aligning with the dynamic nature of language interaction in real contexts(Makransky & Petersen, 2021). Imagination allows learners to transcend existing teaching scenarios and enter cultural contexts and intercultural communication situations that traditional classrooms cannot provide, promoting creative language practice and contextualized expression(Petersen et al., 2022). This learning model has already shown significant advantages in STEM education, for instance, enhancing the visualization of complex concepts and allowing safe exploration through immersive virtual laboratories. This likewise demonstrates its potential to improve pragmatic complexity and cognitive engagement in language learning.
On this basis, metaverse technology further extends the value of immersive language learning by creating shared immersive learning spaces(Wang et al., 2022). Metaverse classrooms can support real-time collaboration, role-playing, cross-cultural simulations, and interactive discussions, providing dynamic social contexts and sustainable virtual communities for language learning(Şalçini & Yerlikaya, 2022). However, the current education system still faces challenges in integrating immersive technologies, such as scalability, device accessibility, network performance, and the professional development of teachers. At the same time, to meet the diverse needs of learners, metaverse English classrooms urgently need to develop design frameworks that integrate technical, educational, and social aspects, allowing virtual learning environments to meet curriculum standards, support personalized learning, and ensure equitable accessibility(Li & Yu, 2023).
Therefore, although immersive language learning has built a solid foundation, literature indicates that the key to future development lies in constructing a comprehensive teaching model for metaverse English classrooms(Lee et al., 2024). This model integrates language learning objectives, interactive virtual scenarios, cross-cultural communication tasks, and intelligent learning analytics to achieve collaborative innovation across technical, social, and educational dimensions. This direction not only responds to the trend of deeply empowering language education through immersive technologies but also provides an important theoretical and practical basis for building a new type of contextualized, community-oriented, and personalized English learning ecosystem.

4. The Multi-Layer Immersive Language Learning Environment Framework

The development of the MILLE framework is grounded in established educational and technological theories, including constructivism, situational cognition, and immersive learning. These theories emphasize active knowledge construction, context-based learning, and experiential engagement, which collectively inform the design of immersive language learning environments.In addition, the identification of the seven core components is informed by prior research on technology-enhanced learning systems, which highlights the importance of infrastructure support, interaction design, personalization, collaboration, assessment, and system management in effective learning environments. By synthesizing these theoretical perspectives and practical considerations, the MILLE framework is constructed as an integrated model that addresses both pedagogical and technological requirements in metaverse-based English learning.

4.1 Framework Design Principles and Structure

The MILLE framework is developed based on several key design principles to ensure the effective integration of pedagogy, technology, and learner support in a metaverse-based English learning environment. First, the principle of integration emphasizes the coordination of multiple system components rather than isolated functionalities. Second, the principle of learner-centeredness highlights the importance of addressing individual differences through adaptive and personalized learning. Third, the principle of interactivity and immersion focuses on enhancing authentic engagement through interactive and context-rich environments. Finally, the principle of sustainability ensures that the framework can be effectively maintained through continuous support, management, and protection mechanisms.

4.2 The MILLE Framework

Guided by these principles, the framework is structured around seven core components, each addressing a specific functional dimension of the learning environment: (As shown in Figure 2.)

◆Infrastructure provides the foundational technological environment for the framework. It includes hardware and network support that enable immersive experiences and ensure the stable operation of virtual learning spaces.

◆Content and interaction focus on the design of learning materials and communication mechanisms. This component supports real-time interaction, task-based activities, and meaningful language use through immersive and interactive content.

◆Personalization and accessibility address learner diversity by enabling adaptive learning pathways. Through personalized content delivery and accessible design, the framework ensures that learners with different needs can effectively participate in the learning process.

◆Collaboration and social emphasize the importance of social learning. This component facilitates peer interaction, group collaboration, and community building, which are essential for developing communicative competence in language learning.

◆Assessment and feedback provide mechanisms for evaluating learning outcomes and supporting continuous improvement. It integrates formative assessment and real-time feedback to monitor progress and guide learners throughout the learning process.

◆Security and privacy ensure a safe and trustworthy learning environment. This includes data protection, identity security, and ethical considerations in the use of immersive technologies.

◆Management and support focus on the operational and instructional aspects of the system. It includes teacher support, resource management, and system maintenance, ensuring the sustainability and effectiveness of the learning environment.

Together, these components form an interconnected system in which technological infrastructure, pedagogical design, and learner support mechanisms operate in coordination. This integrated structure enables the effective design and implementation of metaverse-based English learning environments.

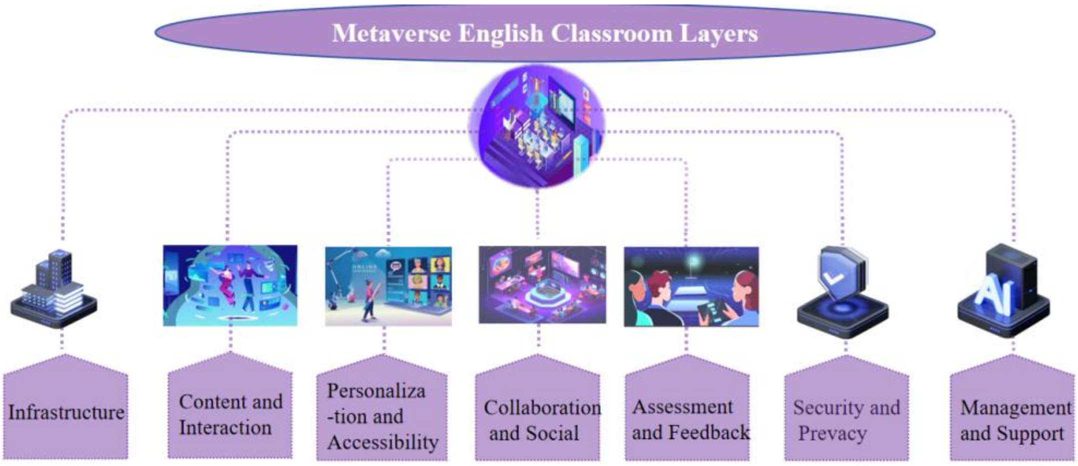

Figure 4.1: The MILLE Framework

4.2.1 Infrastructure

This section will detail how the infrastructure layer lays a solid foundation for higher-level teaching interactions and personalized services from three key dimensions: teaching platform technology, network and connectivity, and hardware compatibility. Its structural diagram is shown in Figure 3.

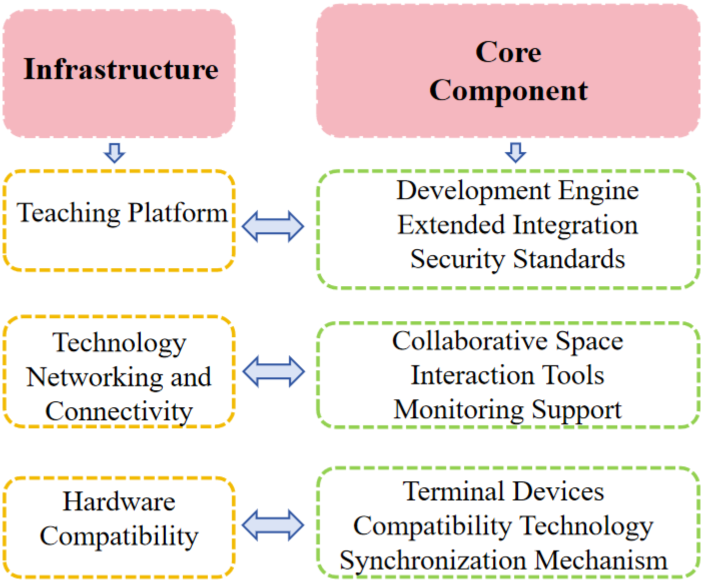

Figure 4.2 Infrastructure

1) Teaching Platform Technology

The platform technology of the metaverse English classroom forms the core foundation for creating an immersive language learning environment, providing students with interactive, multi-scenario virtual learning spaces(Muthmainnah et al., 2023). The platform supports rich three-dimensional spaces and interactive features, making the virtual language learning experience both realistic and manageable. Platforms such as Unity, Unreal Engine, Roblox Education, and Engage XR offer powerful tools for creating highly customized English learning scenarios, with high-quality graphics rendering, virtual character performance, and flexible interaction mechanisms(Lee et al., 2025). Among them, Unity and Unreal Engine are particularly suitable for building multimodal language learning scenarios, such as airports, hotels, international conference rooms, or business negotiation rooms. Their strong rendering capabilities and flexible scripting support enable dialogue triggers, action animations, and virtual character interactions(Yuan et al., 2023). Developers can use these platforms to set up targeted language tasks and interactive resources, allowing teachers to design communicative activities that closely mirror real-life language use.

To achieve an immersive language communication experience, the metaverse English classroom platform must integrate virtual reality (VR) and augmented reality (AR) technologies. VR allows students to deeply engage in simulated conversational scenarios, role-playing, and cross-cultural exchanges, enhancing spoken output and pragmatic skills(İbili et al., 2024). For instance, students can debate in a virtual United Nations meeting room or practice travel English in an airport scenario. AR, on the other hand, can overlay language prompts, vocabulary scaffolds, pronunciation feedback, or cultural annotations in real-world environments. Using tools like ARKit and ARCore, students can explore holographic sentence structures, idiom explanations, or pronunciation examples with mobile devices, enabling blended learning and real-time language support. The combination of VR and AR enhances immersion while providing students with immediate language assistance(AlSaleem, 2023). Scalability is a key feature of the metaverse English classroom platform, ensuring that it can be smoothly accessed by different numbers of users across various devices. The platform should support multiple terminals, including VR headsets, computers, tablets, and smartphones, to reduce the digital divide and improve learning equity. At the same time, the platform needs to efficiently manage learning activities of different scales, whether it's small group discussions, task-based dialogues, or large lectures and simulated meetings, maintaining stable operation(Wei, 2023). For example, Engage XR can support large-scale, cross-regional user interactions, while Unity and Unreal Engine can optimize performance for different devices. Good scalability ensures that the classroom evolves smoothly as student needs grow. In addition, the platform's scalability is reflected in its support for third-party tools and plugins to enhance personalized learning and functionality. For instance, it can integrate AI speech assessment, voice recognition, automatic grammar correction, and learning data analytics tools; it can also connect to platforms like Kahoot or ClassIn through plugins or APIs for vocabulary tests, task grading, and interactive rewards. This scalability allows the classroom to provide personalized support, monitor students' speaking and writing performance in real time, and stimulate learning motivation through interactive assessments. Data security and encryption are crucial in the platform technology, aimed at protecting students' identity information, voice recordings, and learning data(Lee, 2023). The platform must comply with international education security standards such as GDPR and FERPA to ensure the secure storage and use of language interaction logs and biometric voice data. Unity and Unreal provide documentation for data encryption and communication encryption protocols, and platforms like Engage XR also have built-in user privacy protection features. Strong security measures prevent unauthorized access, allowing students to learn safely in virtual communication activities. By prioritizing data security, platform technology provides a stable and reliable environment for immersive English learning.

2) Network and Connectivity

In the metaverse English classroom, the network and connectivity layer provides core technological support for the construction of virtual collaborative spaces, enabling learners to engage in language interaction, cooperative learning, and task-driven communication in immersive contexts, thereby simulating real-life language use scenarios(Ali et al., 2023). These collaborative spaces offer students platforms for group discussions, situational task collaboration, and cross-cultural communication, facilitating the transfer and application of language knowledge in authentic pragmatic contexts(Pulgar et al., 2023). For example, teachers can create diverse virtual scenarios based on teaching tasks (such as airport information halls, international conference rooms, or business negotiation spaces) and organize students into group discussion areas to conduct role-playing, information exchange, and language expression activities. This task-oriented collaborative model not only enhances learning engagement but also allows teachers to provide precise guidance within group spaces, thereby improving learning efficiency and the quality of language output. Shared language tools and collaborative whiteboards further strengthen the quality of interaction among students(Joshi et al., 2024). Learners can collaboratively record vocabulary, syntactic structures, discourse points, or argument logic on virtual whiteboards and continuously refine their language expression through real-time editing. Taking English speech training as an example, students can use collaborative whiteboards to build speech outlines, annotate logical connectors, and revise syntactic expressions, thereby improving their discourse organization and expressive abilities. In addition, these collaborative tools support interoperability across VR headsets, mobile devices, and desktop terminals, providing convenience for students with different technological conditions and learning preferences. Voice and text communication tools are important supports for creating an immersive language learning experience. Voice call functions can simulate the natural rhythm of language communication, allowing students to engage in real-time conversations, debates, or situational exercises, closely mimicking real interactions. At the same time, text chat tools provide support for learners who need more time to organize their language or prefer to record their expressions(Greenan, 2021). Built-in features such as translation and speech recognition can also help students with lower language proficiency improve comprehension and pronunciation, reduce communication anxiety, and enhance learning inclusivity(Chessa & Solari, 2021). For example, in cross-cultural communication simulations, students can leave text messages to record key information while simultaneously participating in voice discussions, achieving a more realistic and efficient interactive experience. Furthermore, document sharing and collaborative editing functions effectively enhance the quality of student collaboration in English writing and project-based tasks. Learners can collaboratively write research reports, organize survey data, design presentation materials, or revise texts within the metaverse platform. For instance, in an English poster creation task, students can discuss topics in a virtual meeting space while simultaneously working on image and text design, language polishing, and content supplementation in shared documents, achieving multimodal expression and collaborative creation. This process not only strengthens language application skills but also promotes the cross-cultural collaboration skills and communication literacy required in the digital age. Finally, teacher monitoring and support tools play a key role at the network and connectivity level(Zhang et al., 2021). Teachers can flexibly enter different collaborative spaces to observe students' language interactions, provide targeted language feedback, and correct expression errors in real time, thereby improving semantic accuracy and expression strategies. For example, during situational debate activities, teachers can enter each group's discussion area to guide students in using more complex argumentative sentences, logical connectors, and polite expression strategies. This precise support not only enhances learning outcomes but also ensures that students with varying language backgrounds can all access effective deep learning opportunities.

3) Hardware Adaptation

In the metaverse English classroom, hardware integration is a key factor in ensuring learners can participate fairly and interact fully(Sanjith, 2024). It determines whether different students can access the virtual language learning environment equally and gain an immersive experience. To achieve technological inclusivity, the metaverse English classroom needs to support multi-device access, including high-end VR headsets, desktops, tablets, and smartphones. Although VR headsets offer the highest level of immersion, multi-platform compatibility ensures that students from different socio-economic backgrounds have the opportunity to participate(Chiriac, 2025).

For example, students without VR devices can still interact in class via computers or mobile devices, preventing learning inequalities caused by hardware differences. High-end VR equipment (such as the Meta Quest, HTC Vive, etc.) can provide highly immersive communicative scenarios for language learning, helping students naturally practice English through interactive experiences(Donally, 2021). For instance, students wearing VR headsets can be placed in simulated environments such as “ordering in a restaurant,” “airport security check,” or “business negotiation,” where they can engage in real-time conversations with virtual characters or peers and learn authentic language usage. However, to avoid tying learning to expensive devices, the metaverse English classroom should also support mid-range and low-end VR devices, enhancing cost accessibility and educational inclusiveness to ensure all learners can access immersive language learning technologies. At the same time, compatibility with desktops and laptops is particularly important for students with limited hardware resources. Using a keyboard and mouse to interact with scenarios, students can still complete role-playing, collaborative virtual tasks, or cross-cultural exchange activities. For example, without wearing a VR headset, students can conduct English interviews in a “virtual press room,” participate in English debates in an “international conference room,” or practice bargaining in English at a “virtual market.” Since most schools are already equipped with traditional computer devices, this compatibility helps implement metaverse English teaching at a lower cost. Meanwhile, the inclusion of tablets and smartphones further expands the possibilities for mobile learning, offering students portable learning options. Tablets, with their larger screens and touch controls, support English speech practice, discourse construction, and virtual task operations, whereas smartphones provide a low-threshold option for students lacking computers or VR devices, particularly useful in areas with higher mobile learning adoption(Gao et al., 2021). Students can even use mobile devices for speaking assessments, task feedback, or group collaboration while on the bus, in the library, or during extracurricular activities, breaking the spatial constraints of traditional classrooms(Bacca-Acosta et al., 2022). Cross-platform synchronization is crucial for ensuring continuous learning. Students can experience immersive oral tasks at school using VR devices, and continue writing, reading, or discourse construction activities at home on a tablet or computer without losing progress or resources. This seamless cross-device mechanism ensures consistency in the learning experience, allowing students to choose learning terminals flexibly based on their environment and needs, thereby creating a metaverse English classroom that is fairer, more flexible, and inclusive.

4.2.2 Content and Interaction Layer

1) 3D Virtual English Learning Environment

3D virtual environments provide authentic and dynamic contextual scaffolding for language acquisition. Relying on high-performance engines such as Unity and Unreal Engine, as well as distributed virtual technologies, this level constructs visualized, story-driven immersive spaces that allow students to enter the same domain with virtual avatars to collaborate and perform tasks(Chen et al., 2021). To meet diversified teaching needs, virtual environment designs usually focus on four core thematic scenarios:

International Student Center: Simulates areas such as consultation desks and course centers, training survival English related to inquiries and freshman adaptation. Global News Studio: Provides anchor desks and interview areas, where students can practice public speaking and information organization through simulated news broadcasts and commentary. International Cultural Exchange Hall: Features multimedia exhibition booths and cultural interaction zones, enhancing intercultural awareness through guided tours and dialogue tasks. Themed Learning Café: Creates low-anxiety semi-realistic life scenarios, helping students build confidence and fluency in speaking through ordering food and casual conversations. These highly realistic virtual settings effectively connect textbook content with real life, enabling students to achieve a natural integration of language input and output through context-driven tasks. As shown in Figure 4.

Figure 4. 3: D Virtual English Learning Environment

2) Interactive Learning Tools

In the metaverse English classroom, a variety of interactive learning tools work together to create an immersive and contextualized language learning environment("Contextual Language Learning," 2021). Firstly, the virtual avatar interaction system allows students to participate in class activities through customized virtual characters and perform actions such as moving, raising hands, and conversing using a keyboard, VR controllers, or other methods. Real speech can be mapped to avatar dialogue, reducing speaking anxiety and enhancing engagement. Secondly, the virtual scenario task board releases tasks, displays hints and clues within contextualized scenarios. Students can click to view task requirements, submit audio or text content, and receive instant system feedback, providing structured support for task-based teaching. Meanwhile, the real-time voice and subtitle system transcribes students' spoken language into subtitles and, in optional modes, provides immediate correction for grammar, pronunciation, and vocabulary usage, helping to improve spoken accuracy and listening comprehension(Lee et al., 2023). The virtual whiteboard and shared screen tools can be used in class for explanations and collaboration, supporting teachers in presenting material and allowing students to co-edit text or draw mind maps, achieving content visualization and collaborative learning. To reinforce experiential learning, the platform also offers a virtual object operation system, enabling students to click, drag, or grab items such as menus, plane tickets, or cultural artifacts to complete tasks, naturally integrating language expression with problem-solving processes(Cantone et al., 2023). Support for group learning comes from virtual group collaboration rooms, where students can discuss, record, and present in separate rooms, with teachers able to join at any time to provide guidance, facilitating project-based and inquiry-based learning. Finally, the AI-driven NPC dialogue system offers students controllable, repeatable, and stress-free speaking practice partners, such as restaurant servers or interviewees, and generates interaction reports to help students improve fluency and language strategies. These tools work collaboratively within the metaverse environment to build an immersive, open, and highly interactive English learning ecosystem.

3) Gamification Elements

The metaverse English classroom extensively integrates gamified elements, creating an immersive language learning experience while effectively enhancing student engagement and sustained involvement(Wulantari et al., 2023). The classroom usually employs a point system to reward students for completing dialogue tasks, participating in interactions, or achieving learning objectives, which can be used to exchange for virtual items or unlock new features. The level and experience point system helps students ‘level up’ their learning as they progress through tasks, forming a visual growth path, while badges and achievement systems are used to record and showcase students’ outstanding performance in speaking, collaborative inquiry, cultural understanding, and more, enhancing their sense of accomplishment. In task design, the metaverse classroom often uses story-based task chains, integrating language learning into coherent storylines, such as checking in at an airport, conducting interviews in a news studio, or making presentations in a cultural center, allowing students to produce language in authentic contexts. The virtual currency system gives students more autonomy, allowing them to purchase helpful items or unlock themed scenes, further motivating exploration(Odinokaya et al., 2024). Leaderboards promote healthy competition by showcasing learning performance, while team-based modes strengthen communication, collaboration, and problem-solving abilities through multi-person challenge tasks. Some activities introduce countdown mechanisms to increase urgency, encouraging students to improve their language reaction speed and fluency. Additionally, the immediate feedback system reinforces correct behaviors through sound effects, lights, scores, or language prompts, helping students adjust their learning strategies in real-time(Esteves et al., 2023). Hidden clues and “easter eggs” within scenarios also encourage active exploration, enhancing the classroom’s fun and participation. Overall, these gamified elements interact to create a highly interactive, well-incentivized, and richly experiential metaverse English classroom, effectively promoting the depth and breadth of students’ language practice. Figure 5 shows the two modes of individual and team challenge learning.

Immersive Airport Check-In **Leaderboard Competition**

Figure 4.4: Gamification elements

4.2.3 Personalization and Accessibility Layers

1) AI-Driven Personalization

In the metaverse English classroom, AI technology relies on real-time learner modeling to create highly personalized learning paths for each student. The system can automatically adjust task difficulty, learning materials, and feedback methods based on a student’s language proficiency, learning pace, interests, and classroom interaction performance(Zhao, 2025). For example, AI can dynamically recommend reading content, speaking practice scenarios, and vocabulary review resources tailored to individual needs, and generate detailed speech analysis reports immediately after students complete speaking exercises, including pronunciation errors, intonation patterns, and grammar usage. With this continuously optimized personalized recommendation mechanism, the metaverse platform can provide precise support for different learners in immersive learning environments, effectively promoting targeted improvements in language skills.

2) Assistive Technology

To ensure that students of different ability levels can equally participate in classroom activities, the metaverse English classroom integrates a variety of assistive technologies. The platform offers real-time captions, speech-to-text, text-to-speech, simplified reading modes, as well as adjustable fonts and interface contrast, helping students with hearing impairments, visual impairments, or reading difficulties smoothly engage in learning(Chen & Wang, 2024). In addition, the immersive interactive system supports various input methods, including keyboard control, touch operations, gesture recognition, and eye-tracking, significantly lowering the operational barrier, allowing students with physical limitations or limited device access to freely move, communicate, and complete tasks. With these thoughtful accessibility designs, the metaverse classroom greatly enhances the inclusivity and accessibility of English learning.

3) Multilingual Support

To meet the needs of learners from different language backgrounds, metaverse English classrooms generally have built-in multilingual support features, including interface language switching, multilingual presentation of task instructions, and AI-based real-time translation tools. When handling complex tasks, students can choose to view prompts in their native language while interacting and expressing themselves in English, which not only reduces the cognitive load but also enhances engagement(Rahmanu & Molnár, 2024). In addition, multilingual systems can facilitate cross-cultural communication among learners in cross-cultural scenario contexts (such as simulated international conferences or cultural exchange spaces), helping them grasp diverse pragmatic norms in authentic contexts. Multilingual support not only improves the convenience of learning but also strengthens the international perspective and openness of metaverse environments, laying a solid technical foundation for conducting globalized English instruction.

4.2.4 Collaboration and Social Layer

1) Avatar Interaction

In the metaverse English classroom, interactions through virtual avatars create a highly immersive and emotionally expressive communicative agent for learners. Students can create personalized virtual avatars with facial expressions, body movements, and posture changes, allowing them to practice nonverbal communication strategies in simulated real-life communicative situations(Kyrlitsias & Michael-Grigoriou, 2022). Compared to the limited presentation through cameras in traditional online classrooms, virtual avatars convey emotions and communicative intentions more authentically, making activities such as role-playing and dialogue simulations more natural and engaging. This enhanced immersive experience not only increases the realism of communicative tasks but also provides students with a more flexible and controlled environment to develop speaking skills, pragmatic competence, and cross-cultural politeness strategies.

2) Virtual Collaboration Space

Virtual collaborative spaces in the metaverse provide highly adaptable learning environments for project-based learning and team collaboration(Orel, 2022). Teachers can create diverse group discussion rooms and project work areas according to task requirements, allowing students to work together in small groups to complete tasks such as designing virtual travel blogs, planning cultural exhibitions, or conducting inter-school joint presentations. In these virtual spaces, students can share resources synchronously, exchange ideas, and collaboratively advance tasks through linguistic negotiation, significantly enhancing collaborative dialogue and deep language processing. The combination of immersive contexts and task-driven learning not only increases learning engagement but also effectively promotes critical thinking, teamwork skills, and the transfer and application of language in authentic situations.

3) Community Participation and Cross-Cultural Communication

The Metaverse English classroom further breaks through the limitations of time and space, allowing students to engage in multi-level community interaction and cross-cultural collaboration with learning communities worldwide(Çelik & Baturay, 2024). By participating in cross-cultural interviews, thematic inquiries, and collaborative story creation projects with other countries or schools, students can deepen their understanding of cultural differences, values, and social norms in authentic cultural contexts. The virtual environment provides a safe, open, and highly interactive communication platform, enabling students to continuously adjust their language strategies in cross-cultural communication, thereby enhancing global awareness and pragmatic skills. Consequently, community participation in the Metaverse scenario becomes not only a vital source of language input and output but also a key way to cultivate students' cross-cultural communication competence.

4.2.5 Assessment and Feedback Layer

1) Scenario-Based Assessment

In the metaverse English classroom, assessment methods have shifted from traditional paper-and-pencil tests to contextualized assessments characterized by authenticity, interactivity, and task orientation. Students demonstrate their language application skills by participating in immersive tasks such as simulated university admission interviews, virtual workplace communication, or customer service dialogues. This type of assessment integrates real communicative needs with dynamic interactive processes, providing a more comprehensive reflection of students' ability to deploy language resources in context, including spoken fluency, pragmatic appropriateness, strategic use, and emotional regulation. Contextualized assessment not only enhances the authenticity and challenge of tasks but also increases students' engagement and sense of value in the assessment process, making assessment itself an extension and facilitator of learning.

2) Real-Time Feedback and Analysis

The AI-driven instant feedback mechanism is an important component of Metaverse English classrooms, providing learners with highly personalized and finely detailed language support. The system can capture students' speech features, phoneme pronunciation, intonation patterns, and usage of grammatical structures in real time, generating corrective suggestions instantly during spoken interactions. For example, AI can provide precise feedback on pronunciation deviations of specific phonemes, improper use of syntactic structures, or deficiencies in discourse organization, helping students quickly correct mistakes in an immersive environment. At the same time, learning analytics tools can visualize students' language development trajectories, enabling teachers to promptly grasp learning trends and implement targeted guidance, thereby significantly improving learning efficiency and feedback quality.

3) Comprehensive Assessment

The metaverse environment provides a multimodal data foundation for implementing comprehensive assessments, allowing evaluation to go beyond a single dimension of language accuracy. The system can integrate students' performance in collaborative tasks, including the quality of teamwork, creative expression, problem-solving strategies, and cross-cultural communication skills, and analyze them in conjunction with language data to form a more multidimensional learning profile(Shin et al., 2025). This comprehensive assessment model emphasizes the joint development of language proficiency, higher-order thinking, and social interaction skills, reflecting the concept of "learning through use" and "assessing during learning." By integrating evidence from multiple dimensions, teachers can provide more constructive feedback, enabling students to achieve overall progress in language, thinking, and cross-cultural literacy. The specific assessment and feedback process is shown in Figure 6.

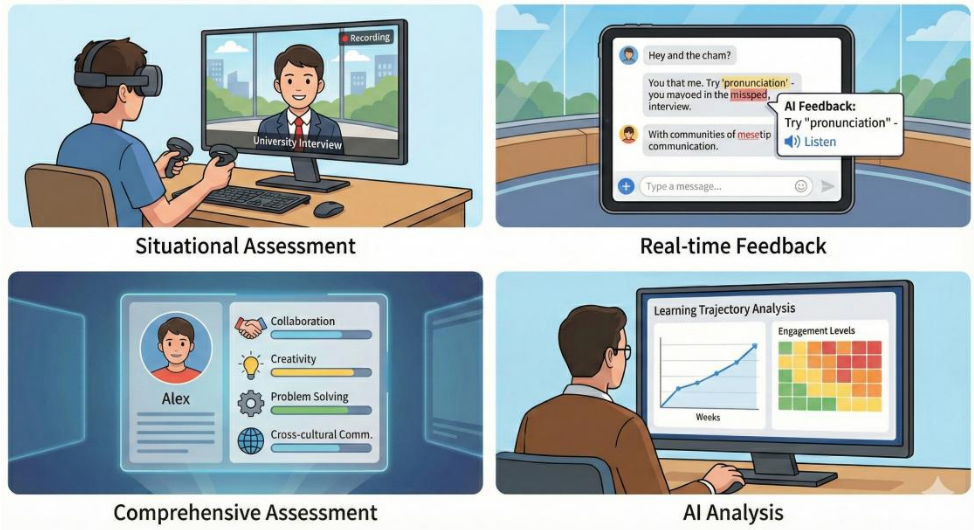

Figure 4.5: Comprehensive Evaluation and Feedback Process

4.2.6 Security and Privacy Layer

1) Data Encryption

In the metaverse English classroom, students' voice data, interaction logs, and learning records are strictly protected using advanced encryption technologies. Whether during data transmission or cloud storage, the system ensures the confidentiality and integrity of information, effectively preventing unauthorized access or data leaks(Jia, 2025). The encryption mechanism not only safeguards students' personal privacy but also allows teachers and educational institutions to conduct data analysis and teaching management legally and securely, enabling the safe use and compliant governance of educational data.

2) Parental Controls

The platform provides parents with comprehensive monitoring and management features to ensure the safety and healthy development of underage students in the virtual environment. Parents can set daily or weekly study time limits, view students' interaction logs, and control access to user-generated content (UGC). Such parental control mechanisms not only prevent excessive use and exposure to inappropriate content but also enhance family supervision and engagement in virtual learning, allowing students to fully participate in immersive learning in a safe and controlled environment.

3) Digital Health

The metaverse English classroom places great emphasis on students' digital health, guiding learners to allocate virtual learning time reasonably through built-in prompts and rest reminders, preventing visual fatigue and the physical and mental strain from prolonged immersion. The system can automatically provide rest suggestions based on students' usage duration and interaction intensity, and guide them through eye relaxation or light activities. This digital health protection mechanism balances the immersive learning experience with physical and mental safety, providing sustainable assurance for students' long-term participation in virtual classrooms(Kim et al., 2025).

4.2.7 Management and Support Layer

1) Teacher Dashboard

The Metaverse English classroom is equipped with a fully functional teacher dashboard that allows teachers to monitor students' learning progress and classroom participation in real time. The dashboard not only provides an overview of overall student progress but also visually presents individual and group interaction activity through the engagement heat map in virtual discussions. In addition, based on AI-generated analytical reports, teachers can quickly identify common language errors, comprehension gaps, or collaboration issues within the class, allowing them to tailor teaching strategies and activity design accordingly. This data-driven management tool significantly enhances the precision of classroom management and teaching effectiveness.

2) Content Management System

The platform's built-in content management system provides teachers with a wealth of virtual scenarios, task templates, and course resources, while supporting highly customizable instructional design. Teachers can flexibly adjust scene elements, interactive tasks, and assessment activities according to course objectives, student abilities, and preferences, enabling personalized and context-rich course delivery. This system not only effectively saves preparation time but also ensures the diversity of virtual learning environments and the relevance of teaching activities, further enhancing student engagement and learning experience in immersive contexts(Rudnik, 2023).

3) Technical Support

To ensure the smooth operation of the Metaverse classroom, the platform provides comprehensive technical support, including in-world help desks, user guides, and real-time online consultations. Teachers and students encountering issues with avatars, interactive scenarios, or collaboration tools can receive immediate guidance and solutions. The embedded technical support system not only lowers the operational barrier but also ensures the continuity of teaching activities and the smoothness of the immersive experience, providing reliable support for efficient interaction between teachers and students in the virtual environment(Hwang et al., 2025).

Together, these components form an interconnected system in which technological infrastructure, pedagogical design, and learner support mechanisms operate in coordination. This integrated structure enables the effective design and implementation of metaverse-based English learning environments.

4.3 Innovative Methods for Building an English Classroom in the Metaverse

As illustrated in Figure 7, the MILLE framework introduces a set of integrated innovations for constructing a metaverse-based English classroom. Rather than focusing on isolated improvements, these innovations collectively form a coherent and multidimensional system. At the pedagogical level, the framework enhances teaching paradigms by emphasizing immersive scenarios and embodied participation, enabling learners to engage in more authentic and context-rich language use. At the technological level, it incorporates advanced interaction methods and intelligent support, facilitating more dynamic and responsive learning experiences. From the learner perspective, the framework enables adaptive and personalized learning through dynamic modeling and multidimensional support mechanisms. In terms of evaluation, it introduces a data-driven assessment approach that shifts from traditional testing to competency-oriented and process-based evaluation. At the system level, the framework is further supported by integrated and protective mechanisms, ensuring seamless learning experiences, inclusivity, and a safe learning environment.

Overall, these innovations transform the metaverse classroom into an integrated ecosystem that connects pedagogy, technology, learner development, and evaluation within a unified framework.

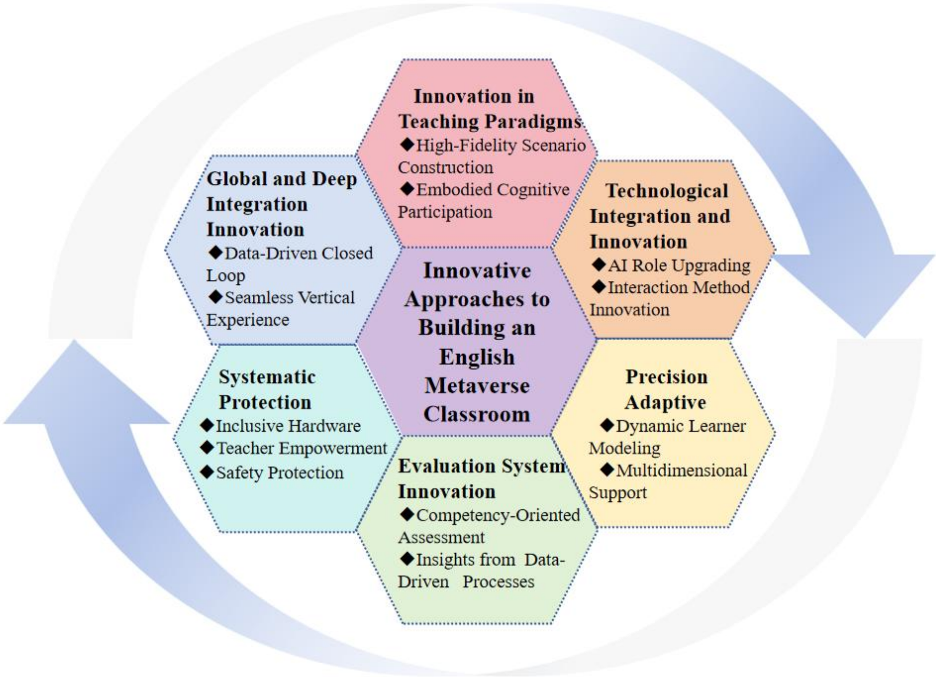

Figure 4.6: Innovative Approaches to Building an English Metaverse Classroom

5. Research Methodology

5.1 Research Design

This study adopts a quantitative research design to provide preliminary empirical support for the proposed MILLE framework. Given that the primary objective of this research is to develop a theoretically grounded framework, the empirical component is designed to evaluate the internal consistency of the measurement model and to examine whether the observed data align with the conceptual structure of the framework.

questionnaire-based survey was employed as the main data collection method. This approach enables the systematic assessment of users' perceptions of key constructs within the MILLE framework, including perceived usefulness, perceived ease of use, immersion, interaction, learning motivation, and behavioral intention. The use of a structured survey instrument allows for statistical analysis of construct relationships and supports the evaluation of measurement reliability and validity.The research design focuses on examining the feasibility and structural rationality of the proposed framework through perception-based data.

5.2 Instrument and Measurement

The measurement instrument used in this study was a structured questionnaire designed to capture users' perceptions of the MILLE framework across multiple dimensions. Based on the theoretical foundations of technology acceptance and immersive learning, six constructs were identified: perceived usefulness (PU), perceived ease of use (PEOU), immersion (IM), interaction (IN), learning motivation (LM), and behavioral intention (BI).

Each construct was measured using multiple items adapted and refined to fit the context of metaverse-based English learning. All items were rated on a five-point Likert scale ranging from

1 (strongly disagree) to 5 (strongly agree). The use of a multi-item measurement approach enhances the reliability and validity of the instrument.

This design ensures both clarity in the main text and completeness of the measurement information. To ensure the quality of the measurement instrument, reliability and validity analyses were conducted. Internal consistency was assessed using Cronbach's alpha, while construct validity was evaluated through exploratory factor analysis. These procedures provide a systematic basis for examining whether the measurement model is consistent with the theoretical structure of the MILLE framework.

5.3 Sample Collection and Data Analysis

Data for this study were collected through a structured questionnaire designed to evaluate users' perceptions of the proposed MILLE framework.

A total of 120 valid responses were obtained, including 85 student participants and 35 teachers. The inclusion of both groups was intended to capture diverse perspectives from key stakeholders in metaverse-based English learning environments. All participants were adults aged 18 years or above, and their participation complied with relevant ethical standards. Participants were selected based on their prior experience with digital or technology-enhanced learning, ensuring that they were capable of providing informed evaluations.

All questionnaires were completed voluntarily and anonymously. Prior to analysis, the data were screened to ensure completeness and consistency. No missing values or extreme outliers were identified, indicating that the dataset was suitable for further statistical analysis. The specific constructs in the questionnaire are as shown in Table 1.

Table 5.1: The Specific Constructs

Construct	Code
Perceived Usefulness	PU1
(PU)	PU2
	PU3
	PU4
Perceived Ease of Use	PEOU1
(PEOU)	PEOU2
	PEOU3
Immersion	IM1
(IM)	IM2
	IM3
Interaction	IN1
(IN)	IN2
	IN3
Learning Motivation	LM1
(LM)	LM2
	LM3
Behavioral Intention	BI1

(BI)	BI2
	BI3

Data analysis was conducted using SPSS. First, descriptive statistics, including mean and standard deviation, were calculated to examine overall user perceptions of the MILLE framework. Second, reliability analysis was performed using Cronbach's alpha to assess the internal consistency of the measurement instrument. Third, exploratory factor analysis (EFA) was conducted to evaluate construct validity and to examine whether the empirical data aligned with the theoretical structure of the framework.

It should be noted that the purpose of the empirical analysis is to provide preliminary support for the reliability of the measurement instrument and the structural rationality of the proposed framework, rather than to assess causal relationships or actual learning outcomes.

6. Research Results

6.1 Users' Perceptions of the MILLE Framework

Descriptive statistics of the questionnaire on cognition and attitudes toward the immersive English learning environment in the metaverse was conducted to examine users' perceptions of the MILLE framework across key constructs, including perceived usefulness, immersion, and interaction. As shown in Table 2.

Table 6.1: Descriptive Statistics of the Questionnaire

Construct	N	Mean	SD	Min	Max
PU	120	4.23	0.62	2	5
PEOU	120	3.95	0.65	2	5
IM	120	4.14	0.66	2	5
IN	120	4.01	0.65	2	5
LM	120	4.22	0.61	3	5

Note: N = 120. All items were measured on a five-point Likert scale (1 = strongly disagree, 5 = strongly agree).

The descriptive statistics indicate that all constructs received relatively high ratings. Behavioral intention achieved the highest mean score (M = 4.34, SD = 0.61), followed by perceived usefulness (M = 4.23, SD = 0.62) and learning motivation (M = 4.22, SD = 0.61), suggesting strong user acceptance of the MILLE framework.

Immersion (M = 4.14, SD = 0.66) and interaction (M = 4.01, SD = 0.65) were also positively evaluated, indicating that the framework provides engaging and interactive learning experiences. Perceived ease of use showed a relatively lower mean (M = 3.95, SD = 0.65), but still reflects general agreement among participants.

Overall, the results suggest that users hold favorable perceptions of the MILLE framework across all measured dimensions.

6.2 Reliability and Validity Assessment

Reliability and validity analyses were conducted to evaluate the quality of the measurement instrument and to examine whether the empirical data align with the theoretical structure of the proposed MILLE framework.

The results indicate that all constructs demonstrate satisfactory internal consistency, with Cronbach's alpha values ranging from 0.84 to 0.91. Specifically, perceived usefulness ($\alpha = 0.89$), perceived ease of use ($\alpha = 0.85$), immersion ($\alpha = 0.87$), interaction ($\alpha = 0.84$), learning motivation ($\alpha = 0.88$), and behavioral intention ($\alpha = 0.91$) all exceed the recommended threshold of 0.70. The reliability test of this questionnaire is as shown in Table 3.

Table 6.2: Reliability Test

Construct	Number of Items	Cronbach's α
PU	4	0.89
PEOU	3	0.85
IM	3	0.87
IN	3	0.84
LM	3	0.88
BI	3	0.91

Construct validity was assessed through exploratory factor analysis. The Kaiser–Meyer–Olkin (KMO) value was 0.91, and Bartlett's test of sphericity was significant ($p < 0.001$), indicating that the data were suitable for factor analysis. As shown in **Table 4.**

Table 6.3 KMO Test

Test	Statistic	Value
Kaiser–Meyer–Olkin (KMO)	-	0.91
Bartlett's Test of Sphericity	χ^2	1856.32
	df	153
	Sig.	< 0.001

All items loaded strongly on their respective constructs, with factor loadings ranging from 0.76 to 0.90. The extracted factor structure is consistent with the theoretical design of the MILLE framework. These results confirm the reliability and validity of the measurement model and provide indirect empirical support for the structural rationality of the proposed framework. As shown in Table 5.

Table 6.4: Factor Analysis

Item	PU	PEOU	IM	IN	LM	BI
PU1	0.81					
PU2	0.84					
PU3	0.79					
PU4	0.83					
PEOU1		0.76				
PEOU2		0.81				
PEOU3		0.78				
IM1			0.82			
IM2			0.85			

IM3			0.80			
IN1				0.79		
IN2				0.82		
IN3				0.77		
LM1					0.83	
LM2					0.86	
LM3					0.81	
BI1						0.88
BI2						0.90
BI3						0.87

6.3 Summary of Empirical Support

Overall, the empirical results provide supportive evidence for the proposed MILLE framework. The reliability and validity analyses confirm that the measurement instrument is robust and aligns with the theoretical structure, while the descriptive findings indicate positive user perceptions.

It should be noted that the purpose of this empirical analysis is not to establish causal effects or learning outcomes, but rather to provide preliminary support for the framework's structural rationality and user acceptance.

Therefore, the findings serve as supplementary evidence reinforcing the feasibility and conceptual soundness of the MILLE framework.

7. Discussion

From a measurement perspective, the strong internal consistency and clear factor structure indicate that the instrument effectively captures the key constructs underlying the MILLE framework. The alignment between the extracted factors and the theoretical design suggests that the framework is structurally sound and that its core dimensions are empirically distinguishable. This consistency provides indirect support for the rationality of the framework and reinforces its theoretical foundation.

In terms of user perception, the high mean scores across constructs indicate that participants generally recognize the value of the MILLE framework. In particular, the high ratings for behavioral intention and perceived usefulness suggest that users are willing to adopt the framework and perceive it as beneficial for English learning. The positive evaluations of immersion and interaction further highlight the importance of experiential and communicative elements in metaverse-based learning environments. These findings suggest that integrating immersive and interactive features into language learning frameworks can enhance user engagement and acceptance.

From a theoretical perspective, this study contributes to the growing body of research on technology-enhanced language learning by proposing and empirically supporting a multilayer framework that integrates usability, immersion, and interaction. The findings are consistent with established theories such as the Technology Acceptance Model, in which perceived usefulness plays a central role in shaping behavioral intention. At the same time, the inclusion of immersion and interaction extends traditional models by incorporating experiential dimensions that are particularly relevant in metaverse environments. This highlights the need to move beyond purely functional perspectives and to consider user experience as a core component of learning framework design.

It should be noted that the empirical analysis in this study is based on user perception data rather than objective measures of learning performance. Therefore, the findings should be interpreted as providing preliminary support for the feasibility and structural rationality of the MILLE framework, rather than definitive evidence of its effectiveness in improving learning outcomes. Future research is needed to further validate the framework through experimental designs, longitudinal studies, and performance-based evaluations.

Overall, this study bridges the gap between conceptual framework design and empirical investigation. By combining theoretical modeling with empirical support, it offers a structured and evidence-informed approach to the design of metaverse-based language learning environments.

8. Research Objectives and Significance

This study aims to develop and preliminarily validate the MILLE framework for metaverse-based English learning. Specifically, it seeks to: 1) propose a structured and theoretically grounded model that integrates pedagogical, technological, and systemic components, 2) examine the reliability and validity of the measurement instrument used to operationalize the framework, and 3) explore users' perceptions of the framework in terms of usefulness, immersion, and interaction.

The significance of this study lies in both theoretical and practical contributions. Theoretically, it extends existing research by providing a systematic framework that integrates immersive technologies with established learning theories, addressing the fragmentation commonly observed in prior studies. Practically, the MILLE framework offers structured guidance for the design and implementation of metaverse-based English classrooms, supporting educators and system developers in creating more engaging and adaptive learning environments.

Furthermore, by incorporating preliminary empirical validation, this study bridges the gap between conceptual framework development and data-driven evaluation, contributing to a more robust foundation for future research in immersive language learning.

9. Limitations

Despite its contributions, this study has several limitations. The empirical component is based on a relatively small sample size of 120 participants, which may restrict the generalizability of the findings. Future research could incorporate larger and more diverse samples to enhance external validity.

The study relies on self-reported questionnaire data, which may be subject to response bias. Although reliability and validity analyses were conducted, the inclusion of additional data sources, such as behavioral data or experimental observations, would provide more robust evidence.

Moreover, the empirical analysis primarily focuses on the reliability and validity of the measurement model as well as users' perceptions, rather than examining actual learning outcomes. Consequently, the effectiveness of the MILLE framework in improving language proficiency remains to be further investigated.

10. Conclusion

This study proposes the MILLE framework as a structured model for metaverse-based English learning, integrating pedagogical principles, technological affordances, and system-level support into a unified framework. By organizing key components such as infrastructure, interaction, personalization, collaboration, assessment, and security, the framework provides a comprehensive approach to designing immersive language learning environments.

In addition to its conceptual contribution, this study incorporates an empirical component to provide preliminary validation. The results indicate that the measurement model demonstrates satisfactory reliability and validity, and that users hold generally positive perceptions of the framework in terms of usefulness, immersion, and interaction.

Overall, the findings suggest that the MILLE framework is both theoretically grounded and practically feasible. It offers a systematic foundation for future research and application in metaverse-based English education. Future studies are encouraged to further validate the framework through large-scale empirical research and real-world implementation, thereby advancing the development of immersive and adaptive language learning environments.

References

Boumová, V. (2008). *Traditional vs. modern teaching methods: Advantages and disadvantages of each* [Master's thesis, Masaryk University, Faculty of Arts].

Sharma, N., & Joshi, H. (2024). Traditional vs. modern English language teaching methods: Study based on a survey. *MIER Journal of Educational Studies Trends and Practices*, 21–36.

Nguyen, T. N. N., & Nguyen, T. T. U. (2024). Benefits of teaching English to children in virtual classes: Teachers' perspectives from Khanh Hoa province, Viet Nam. *International Journal of Language Instruction, 3*(1), 91–107.

Efgivia, M. G., Rinanda, R. Y. A., Hidayat, A., et al. (2021). Analysis of constructivism learning theory. In *Proceedings of the 1st UMGESHIC International Seminar on Health, Social Science and Humanities (UMGESHIC-ISHSSH 2020)* (pp. 208–212). Atlantis Press.

Hwang, Y., & Lee, J. Y. (2024). Exploring pre-service English teachers' perceptions and technological acceptance of metaverse language classroom design. *SAGE Open, 14*(4), 21582440241300543.

Kirshner, D., & Whitson, J. A. (2021). Editors' introduction to situated cognition: Social, semiotic, and psychological perspectives. In *Situated cognition* (pp. 1–16). Routledge.

Lui, T. W., & Goel, L. (2022). Learning effectiveness of 3D virtual reality in hospitality training: A situated cognitive perspective. *Journal of Hospitality and Tourism Technology, 13*(3), 441–460.

Rossetti, G. (2023). Applying Kolb's experiential learning theory to an event management course: Practical guidelines for educators. *Event Management, 27*(7), 1025.

Guo, H., & Gao, W. (2022). Metaverse-powered experiential situational English-teaching design: An emotion-based analysis method. *Frontiers in Psychology, 13*, 859159.

Chen, H. J. (2024). Gather in the metaverse: Learning outcomes, virtual presence, and perceptions of high- and low-achieving pre-service teachers of English as a foreign language. *Education and Information Technologies, 29*(7), 8549–8577.

Therrien, C. (2023). Immersion. In *The Routledge companion to video game studies* (pp. 564–571). Routledge.

Frost, S., Kannis-Dymand, L., Schaffer, V., et al. (2022). Virtual immersion in nature and psychological well-being: A systematic literature review. *Journal of Environmental Psychology, 80*, 101765.

Farina, M. (2021). Embodied cognition: Dimensions, domains and applications. *Adaptive Behavior, 29*(1), 73–88.
Yagcioglu, O. (2023). Embodied cognition and foreign language learning in our classes. *European Journal of English Language Teaching, 7*(6).
Wu, Z. (2025). The development and limitations of digital assessment in language testing. In *SHS Web of Conferences, 222*, 01009. EDP Sciences.
Choi, W. C., & Chang, C. I. (2025). Advantages and limitations of open-source versus commercial large language models (LLMs): A comparative study of DeepSeek and OpenAI's ChatGPT.
Parveen, S., Kang, M. B. A., & Khan, M. A. (2025). ECO-ENGLISH: Cultivating language skills through sustainable practices. *Journal of Applied Linguistics and TESOL (JALT), 8*(4), 421–429.
Chiu, T. K. F., Lin, T. J., & Lonka, K. (2021). Motivating online learning: The challenges of COVID-19 and beyond. *The Asia-Pacific Education Researcher, 30*(3), 187–190.
Peng, R., & Fu, R. (2021). The effect of Chinese EFL students' learning motivation on learning outcomes within a blended learning environment. *Australasian Journal of Educational Technology, 37*(6), 61–74.
Bahari, A. (2023). Affordances and challenges of technology-assisted language learning for motivation: A systematic review. *Interactive Learning Environments, 31*(9), 5853–5873.
Van, L. K., Dang, T. A., Pham, D. B. T., et al. (2021). The effectiveness of using technology in learning English. *AsiaCALL Online Journal, 12*(2), 24–40.
Purwantoro, A., Asari, S., & Maruf, N. (2021). The effectiveness of E-Learning madrasah in English teaching and learning. *BIRCI-Journal: Humanities and Social Sciences, 4*(3), 5234–5244.
Ariffin, K., Abdul Halim, N., & Darus, N. A. (2021). Discovering students' strategies in learning English online. *Asian Journal of University Education, 17*(1), 261–268.
Hein, R. M., Wienrich, C., & Latoschik, M. E. (2021). A systematic review of foreign language learning with immersive technologies (2001–2020). *AIMS Electronics and Electrical Engineering, 5*(2), 117–145.
Weng, Y., Schmidt, M., Huang, W., et al. (2024). The effectiveness of immersive learning technologies in K–12 English as second language learning: A systematic review. *ReCALL, 36*(2), 210–229.
Dhimolea, T. K., Kaplan-Rakowski, R., & Lin, L. (2022). A systematic review of research on high-immersion virtual reality for language learning. *TechTrends, 66*(5), 810–824.
Fernandes, F. A., Rodrigues, C. S. C., Teixeira, E. N., et al. (2023). Immersive learning frameworks: A systematic literature review. *IEEE Transactions on Learning Technologies, 16*(5), 736–747.
Beck, D., Morgado, L., & O'Shea, P. (2023). Educational practices and strategies with immersive learning environments: Mapping of reviews for using the metaverse. *IEEE Transactions on Learning Technologies, 17*, 319–341.
Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review, 33*(3), 937–958.
Petersen, G. B., Petkakis, G., & Makransky, G. (2022). A study of how immersion and interactivity drive VR learning. *Computers & Education, 179*, 104429.
Wang, W., Zhou, F., Wan, Y., et al. (2022). A survey of metaverse technology. *Chinese Journal of Engineering, 44*(4), 744–756.
Şalçını, S., & Yerlikaya, T. (2022). Metaverse: Technology of the future. *Prizren Social Science Journal, 6*(3), 55–63.
Li, M., & Yu, Z. (2023). A systematic review on the metaverse-based blended English learning. *Frontiers in Psychology, 13*, 1087508.

Lee, S. M., Yang, Z., & Wu, J. G. (2024). Live, play, and learn: Language learner engagement in the immersive VR environment. *Education and Information Technologies, 29*(9), 10529–10550.
Muthmainnah, M., Khang, A., Al Yakin, A., et al. (2023). An innovative teaching model: The potential of metaverse for English learning. In *Handbook of research on AI-based technologies and applications in the era of the metaverse* (pp. 105–126). IGI Global.
Lee, S., Jeon, J., & Choe, H. (2025). Enhancing pre-service teachers' Global Englishes awareness with technology: A focus on AI chatbots in 3D metaverse environments. *TESOL Quarterly, 59*(1), 49–74.
Yuan, J., Liu, Y., Han, X., et al. (2023). Educational metaverse: An exploration and practice of VR wisdom teaching model in Chinese Open University English course. *Interactive Technology and Smart Education, 20*(3), 403–421.
İbili, E., Ölmez, M., İbili, A. B., et al. (2024). Assessing the effectiveness and student perceptions of synchronous online flipped learning supported by a metaverse-based platform in medical English education: A mixed-methods study. *Education and Information Technologies, 29*(14), 18643–18673.
Issa Ahmad AlSaleem, B. (2023). The efficiency of metaverse platforms in language learning based on Jordanian young learners' perceptions. *Arab World English Journal, 14*.
Wei, Z. (2023). Metaverse-based online English teaching scheme in multi-source and cross-domain environment. *Fractals, 31*(6), 2340153.
Lee, S. M. (2023). Second language learning through an emergent narrative in a narrative-rich customizable metaverse platform. *IEEE Transactions on Learning Technologies, 16*(6), 1071–1081.
Ali, M., Naeem, F., Kaddoum, G., et al. (2023). Metaverse communications, networking, security, and applications: Research issues, state-of-the-art, and future directions. *IEEE Communications Surveys & Tutorials, 26*(2), 1238–1278.
Pulgar, J., Ramírez, D., & Candia, C. (2023). Enhanced social connectivity in hybrid classrooms versus academic centrality in online settings. *Physical Review Physics Education Research, 19*(2), 020155.
Joshi, K., Kumar, R., Bharany, S., et al. (2024). Exploring the connectivity between education 4.0 and classroom 4.0: Technologies, student perspectives, and engagement in the digital era. *IEEE Access, 12*, 24179–24204.
Greenan, K. A. (2021). The influence of virtual education on classroom culture. *Frontiers in Communication, 6*, 641214.
Chessa, M., & Solari, F. (2021). The sense of being there during online classes: Analysis of usability and presence in web-conferencing systems and virtual reality social platforms. *Behaviour & Information Technology, 40*(12), 1237–1249.
Zhang, X., Gao, X., Yi, H., et al. (2021). [Retracted] Design of an intelligent virtual classroom platform for ideological and political education based on the mobile terminal APP mode of the Internet of Things. *Mathematical Problems in Engineering, 2021*, 9914790.
Sanjith, S. L. (2024). Designing a smart classroom with advanced audiovisual equipment and soft collaboration capabilities for higher education. *International Journal, 13*(3), 1–8.
Chiriac, A. L. (2025). *Development of a multi-platform application for the University Virtual Campus*.
Donally, J. (2021). *The immersive classroom: Create customized learning experiences with AR/VR*. ASCD.
Gao, H., Bozkir, E., Hasenbein, L., et al. (2021). Digital transformations of classrooms in virtual reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–10).
Bacca-Acosta, J., Tejada, J., Fabregat, R., et al. (2022). Scaffolding in immersive virtual reality environments for learning English: An eye tracking study. *Educational Technology Research and Development, 70*(1), 339–362.

Chen, C. H., Hung, H. T., & Yeh, H. C. (2021). Virtual reality in problem-based learning contexts: Effects on the problem-solving performance, vocabulary acquisition and motivation of English language learners. *Journal of Computer Assisted Learning, 37*(3), 851–860.
Contextual language learning: Real language learning on the continuum from virtuality to reality. (2021). Springer Nature.
Lee, H., Hsia, C. C., Tsoy, A., et al. (2023). VisionARy: Exploratory research on contextual language learning using AR glasses with ChatGPT. In *Proceedings of the 15th Biannual Conference of the Italian SIGCHI Chapter* (pp. 1–6).
Cantone, A. A., Francese, R., Sais, R., et al. (2023). Contextualized experiential language learning in the metaverse. In *Proceedings of the 15th Biannual Conference of the Italian SIGCHI Chapter* (pp. 1–7).
Wulantari, N. P., Rachman, A., Sari, M. N., et al. (2023). The role of gamification in English language teaching: A literature review. *Journal on Education, 6*(1), 2847–2856.
Odinokaya, M., Krylova, E., Pyatnitsky, A., et al. (2024). Immersive storytelling and gamification elements as a tool for foreign language epideictic speech development. In *International Conference on Professional Culture of the Specialist of the Future* (pp. 162–188). Springer Nature Switzerland.
Esteves, J. R., Cardoso, J. C. S., & Gonçalves, B. S. (2023). Design recommendations for immersive virtual reality application for English learning: A systematic review. *Computers, 12*(11), 236.
Zhao, J. (2025). Advancing English language education: A mixed-methods analysis of AI-driven tools' impact on engagement, personalization, and learning outcomes. *Education and Information Technologies*, 1–41.
Chen, S., & Wang, J. (2024). Virtual reality human–computer interactive English education experience system based on mobile terminal. *International Journal of Human–Computer Interaction, 40*(13), 3560–3569.
Rahmanu, I. W. E. D., & Molnár, G. (2024). Multimodal immersion in English language learning in higher education: A systematic review. *Heliyon, 10*(19).
Kyrlitsias, C., & Michael-Grigoriou, D. (2022). Social interaction with agents and avatars in immersive virtual environments: A survey. *Frontiers in Virtual Reality, 2*, 786665.
Orel, M. (2022). Collaboration potential in virtual reality (VR) office space. In *Transforming the workplace of tomorrow*. Springer.
Çelik, F., & Baturay, M. H. (2024). The effect of metaverse on L2 vocabulary learning, retention, student engagement, presence, and community feeling. *BMC Psychology, 12*(1), 58.
Shin, D., Kwon, S. K., Noh, W. I., et al. (2025). Exploring the role of the metaverse in English speaking proficiency tests. *Journal of Computer Assisted Learning, 41*(1), e13108.
Jia, L. (2025). Research on English classroom data security protection technology based on intelligent algorithm. *International Journal of High Speed Electronics and Systems*, 2540465.
Kim, H., Kang, J. H., Kwon, Y., et al. (2025). Development of a smart health care service using metaverse and chatbot technologies for adolescents, parents, and school health teachers: User-centered design approach. *Journal of Medical Internet Research, 27*, e69190.
Руднік, Ю. В. (2023). The use of the metaverse studio application in foreign language teaching as an innovative AR technology. *Відкрите освітнє e-середовище сучасного університету*, (14), 40–49.
Hwang, Y., Lee, S., & Jeon, J. (2025). Integrating AI chatbots into the metaverse: Pre-service English teachers' design works and perceptions. *Education and Information Technologies, 30*(4), 4099–4130.