This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

TLT-2026-01-0059

Designing Immersive Multi-User Virtual Environments for Collaborative Learning: A Multidimensional Framework and Evaluation

Yueyao Qu, Yun Zhou, Jin Shu, Junjun Zhai, Jie Bai, Jialin Yu, Tao Xu**

***Abstract*—Multi-User Virtual Environments (MUVEs) have gained increasing attention in collaborative learning research due to their interactive affordances. However, most current MUVEs remain desktop-based, limiting their ability to simulate authentic learning contexts. Immersive virtual reality (IVR), delivered through head-mounted displays that engage learners' peripheral vision, can enhance their sense of presence and immersion within virtual environments. To explore this potential, the present study proposes a multidimensional design framework that encompasses social, motivational, cognitive, and experiential dimensions and uses it to develop an Immersive Multi-User Virtual Environment (MUIVE). The effectiveness of MUIVE was evaluated using measures of learning performance, social presence and collaboration quality (social), learning motivation (motivational), cognitive load (cognitive), and IVR experience, trust, ease of use, and satisfaction (experiential). Results showed a higher overall performance gain for MUIVE in the gain-score analysis, although the corresponding baseline-adjusted analysis did not yield a statistically significant condition effect. MUIVE significantly enhanced social presence, learning motivation, trust, ease of use, and satisfaction, without increasing cognitive load. Based on these findings, we propose five core design principles for building effective MUIVEs: (a) face-to-face authenticity, (b) engagement enhancement, (c) extraneous cognitive load minimization, (d) avatar embodiment, and (e) system experience. Furthermore, peer familiarity moderated the effect of platform on transfer performance, with MUIVE showing a stronger advantage among unfamiliar dyads. These findings offer design implications for the development of immersive collaborative learning systems and serve as a foundation for future research on optimizing multi-user interaction and learning support in IVR-based environments.**

***Index Terms*—Collaborative learning, empirical design and evaluation, immersive virtual reality (IVR), MUVE**

Data availability statement: The data are available from the corresponding author upon reasonable request.

Funding statement: This work was supported by the Research Projects of the National Natural Science Foundation of China under Grant [62377039]; 2025 Shaanxi Normal University Graduate Education Comprehensive Reform Research and Practice Project under Grant [GERP-25-01]; and Higher Education Research Fund of Northwestern Polytechnical University in 2025, No. GJJJM202504.*(Corresponding author: Yun Zhou,Tao Xu,Yueyao Qu,, Jin Shu, Junjun Zhai, Jie Bai, Jialin Yu)*Conflict of interest statement: The authors declare no conflicts of interest.

Yueyao Qu is with the Faculty of Education, Shaanxi Normal University, Xi'an 710062, Shaanxi Province, P.R. China (e-mail: quyueyao@snnu.edu.cn)

Yun Zhou is with the Faculty of Education, Shaanxi Normal University, Xi'an 710062, Shaanxi Province, P.R. China (e-mail: zhouyun@snnu.edu.cn)

Jin Shu is with the Faculty of Education, Shaanxi Normal University, Xi' an 710062, Shaanxi Province, P.R. China (e-mail: shujin@snnu.edu.cn)

Junjun Zhai is with the Faculty of Education, Shaanxi Normal University, Xi'an 710062, Shaanxi Province, P.R. China (e-mail: zhaijunjun@snnu.edu.cn)

Jie Bai is with the Faculty of Education, Shaanxi Normal University, Xi'an 710062, Shaanxi Province, P.R. China (e-mail: baijie@snnu.edu.cn)

Jialin Yu is with the Faculty of Education, Shaanxi Normal University, Xi' an 710062, Shaanxi Province, P.R. China (e-mail: yujialin@snnu.edu.cn)

Tao Xu is with the School of Software, Northwestern Polytechnical University, Xi ' an 710072, Shaanxi Province, P.R. China (e-mail: xutao@nwpu.edu.cn).

Consent to participate: All participants signed informed consent agreement.

Artificial intelligence generated content: None.

Ethics approval statement: This study was approved by the Ethics Committee of the Faculty of Education at Shaanxi Normal University (No. 2024-161). The study complies with the WMA Declaration of Helsinki regarding participant recruitment, data collection, data use, data confidentiality, and protection of participants' rights. All participants were informed of the voluntary nature of their participation and provided written informed consent. They were assured of anonymity and informed about the purpose of the research, how their data would be used, and that no risks were associated with their involvement. Among the 104 participants, only one was a minor (17 years old), for whom guardian consent was obtained.

I. INTRODUCTION

Multi-User Virtual Environments (MUVEs), which facilitate real-time avatar-based interaction in shared virtual spaces, offer enhanced technological support for collaborative learning and foster a stronger sense of presence and immersion than text-based online chat platforms such as Internet Relay Chat [1]. Beyond synchronous interaction, MUVEs can also incorporate intelligent agents, discussion tools, and diverse digital resources, thereby providing a richer and more flexible collaborative learning experience [2]. Platforms such as Second Life exemplify the application of MUVEs in educational contexts and have been widely adopted in related research [3]. A growing body of empirical studies has demonstrated their effectiveness in enhancing motivation, engagement, academic performance, language acquisition, and collaboration quality (e.g., [4], [5]).

However, most existing MUVE platforms primarily rely on desktop-based virtual systems. Immersive virtual reality (IVR), delivered through head-mounted displays (HMDs) that engage learners' peripheral vision, provides a heightened sense of realism and interactivity [6]. IVR can also evoke authentic physical and emotional responses, such as sensations of weightlessness and acrophobia, which are rarely achievable in desktop-based systems [7]. Emerging research suggests that such immersive experiences can enhance motivation and, in turn, improve learning outcomes [8]. These findings highlight the importance of integrating MUVEs into IVR as a critical direction for future research. Yet, the majority of MUVE studies still focus on desktop-based environments, with limited empirical work and few well-established frameworks

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

addressing the design of immersive MUVEs [9].

To address these limitations, this study designed and developed an Immersive Multi-User Virtual Environment (MUIVE) to support collaborative learning within IVR. MUIVE aims to deliver a more engaging and interactive experience than conventional platforms (e.g., Tencent Meeting), thereby enhancing peer collaboration and learning. To explore this potential, the present study proposes a multidimensional design framework encompassing four dimensions, including social, motivational, cognitive, and experiential, to guide the development of MUIVE. The system's effectiveness was evaluated across corresponding subdimensions: social presence and collaboration quality (social), learning motivation (motivational), cognitive load (cognitive), and IVR experience, trust, ease of use, and satisfaction (experiential).

Each dimension informed specific evaluation criteria, which correspond to the following research questions:

RQ1 (performance): To what extent does MUIVE enhance learners' knowledge acquisition compared with the conventional online platform Tencent Meeting?

RQ2 (social): How does MUIVE influence learners' social presence and collaboration quality?

RQ3 (motivational): How does MUIVE influence learners' motivation to engage in collaborative learning?

RQ4 (cognitive): How does MUIVE influence learners' cognitive load compared with Tencent Meeting?

RQ5 (experiential): How do learners perceive MUIVE in terms of embodiment, physical presence, trust, ease of use, and satisfaction?

Based on these findings, we propose five design principles to inform future development of IVR-supported multi-user collaborative learning environments.

II. Background and Related Work

A. Collaborative Learning

Collaborative learning entails learners interacting, cooperating, and communicating to co-construct knowledge and solve problems together [10]. By engaging in perspective-sharing and mutual inquiry, learners not only deepen their understanding but also stimulate new ideas, thereby promoting cognitive development [11]. As Zambrano et al. observed, collaborative learning can be more effective than individual learning in complex tasks, especially when learners have limited prior knowledge [12].

In recent years, and even more so in the aftermath of the COVID-19 pandemic, online collaborative platforms have seen more widespread adoption in higher education. These platforms help overcome geographical constraints. For example, Gaad's implementation of online collaborative learning in physical science education led to significant improvements in students' achievement and engagement [13]. However, such platforms often fall short in replicating the immersive nature of face-to-face interaction, which limits learner engagement, interaction, and emotional connection, thereby compromising the collaborative learning experience [14]. This highlights the pressing need for online environments that can better replicate face-to-face collaboration to maximize peer support and cooperation.

B. MUVEs for Collaborative Learning

MUVEs are virtual spaces that support multiple users to collaborate in real time by navigating, interacting, manipulating virtual objects, and accessing shared resources [1], [2], [15]. Their synchronous and interactive features have positioned them as a promising tool for online and remote collaborative learning contexts. Existing research has examined MUVEs from multiple perspectives, including learning performance, user experience, and system design, particularly within STEM disciplines such as mathematics, physics, computer science, engineering, and information technology [3].

MUVEs have proven effective for supporting both abstract knowledge (e.g., in mathematics [16] and information technology [17]) and phenomenon-based knowledge (e.g., in geography [18] and biology [19]). For abstract knowledge, learners typically engage in collaborative tasks within MUVEs to facilitate peer understanding and knowledge construction [16], [17]. For phenomenon-based knowledge, collaboration often involves observation and discussion of simulated objects [18], [19]. A majority of empirical evidence highlights the benefits of MUVEs in fostering learners' motivation, academic performance, and collaborative experiences [3]. Recent studies have also extended the focus toward MUVE system design, investigating features such as intelligent agents [20], personalized resource delivery [21], and various modes of presenting instructional materials [22].

However, most existing MUVEs remain desktop-based. Due to the technical barriers associated with IVR, research on immersive MUVEs in education remains limited [23]. Although early efforts have begun to explore the integration of MUVEs within IVR to enhance collaborative learning, this line of inquiry is still in its nascent stage.

C. Towards MUVEs in IVR for Collaborative Learning

IVR creates highly engaging learning environments through technologies such as HMDs and CAVE systems [24]. HMDs enhance immersion by providing a wide field of view that more closely matches human visual perception [25]. Prior studies have shown that learners in immersive virtual reality classrooms generally exhibit better concentration and improved learning performance [26]. CAVE systems, projecting visuals onto surrounding walls to create a shared immersive space, have also proven effective for collaborative tasks, such as De Back et al.'s work on brain structure learning [27]. Compared with desktop-based VR, IVR affords higher immersion, presence, and interaction, and has been associated with enhanced emotional engagement and performance [6].

While IVR's benefits for individual learning are well explored, most existing research has focused on single-user systems, leaving multi-user immersive environments underexplored [11]. In recent work, Conesa et al. developed a novel collaborative virtual environment for spatial skills

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

training and argued that virtual collaborative environments, which connect multiple learners to interact in real time within a shared virtual space, have considerable potential for STEM education [9]. Although such efforts mark an important step forward, research on multi-user IVR systems remains in its nascent stage. Further investigation is needed to develop frameworks and practical guidelines that can support effective collaboration in these immersive environments.

D. Existing Design Principles for Multimedia Learning and Immersive Collaborative Learning Environments

Research on multimedia learning has established several influential design principles for promoting meaningful learning from words and graphics. In particular, the cognitive theory of multimedia learning emphasizes reducing extraneous processing, managing essential processing, and fostering generative processing through principles such as coherence, segmenting, pretraining, and, more recently, embodiment and immersion. These principles provide an important basis for designing learning materials and interactions in technology-enhanced environments [28].

As immersive virtual reality has developed, collaborative learning has imposed more specific requirements on system design. Existing studies have therefore proposed principle-oriented insights more directly related to immersive collaborative activity. Paulsen et al. argued that design in IVR should address the environment, the activity, and the technology simultaneously, including participation, shared goals, scaffolds, reflection, usability, realism, and interaction support [29]. De Back et al. highlighted coherence, segmenting, pre-training, natural interaction, clear feedback, and support for collaboration through equal access, active participation, interdependence, and visibility of peer behavior [27]. In TICOL, Makransky and Petersen emphasized technological features, social affordances, and pedagogical techniques as shaping social presence, physical presence, body ownership, and agency [11]. Paulsen and Davidsen stressed situated activity, multi-perspective joint analysis, and authentic reflection [30].

These studies provide an important foundation for design in multimedia, VR, and immersive collaborative contexts. However, research on immersive multi-user virtual environments for collaborative learning remains limited, and a clearly established framework for guiding design in this specific context is still lacking. This gap motivates the present study, which first develops a multidimensional framework for MUIVE design based on prior studies and then, drawing on the empirical findings, proposes a more focused set of design principles for immersive multi-user collaborative learning.

III. Design Rationale and Implementation

A. Design Rationale of MUIVE

Drawing on prior studies, we identified six critical factors in the design and evaluation of multi-user IVR learning systems, including social [11], motivational [5], cognitive load [5], IVR-specific immersive experiences [31], technological experience [29], and user satisfaction [32]. Based on these factors, we proposed a conceptual framework to inform the design and development of MUIVE including four dimensions: social, motivational, cognitive, and experiential (see Fig. 1). These four dimensions constitute the higher-order structure of the framework and represent the key design and learning-process conditions through which MUIVE is expected to support collaborative learning. Rather than evaluating MUIVE only through subjective experience variables, the framework links these multidimensional design factors to both subjective assessments and performance-based outcomes. In this study, learning performance functions as an outcome layer of the framework and is assessed separately through knowledge retention, knowledge transfer, and performance in collaborative tasks. Thus, the framework serves not only to capture learners' experiences in MUIVE, but also to examine whether these design dimensions translate into measurable collaborative learning gains.

The specific variables used in the present study, such as social presence, collaboration quality, trust, ease of use, and satisfaction, represent contextually grounded operationalizations of these broader dimensions. Accordingly, the framework may also accommodate alternative or additional indicators depending on the learning task, disciplinary domain, and immersive collaborative context under investigation.

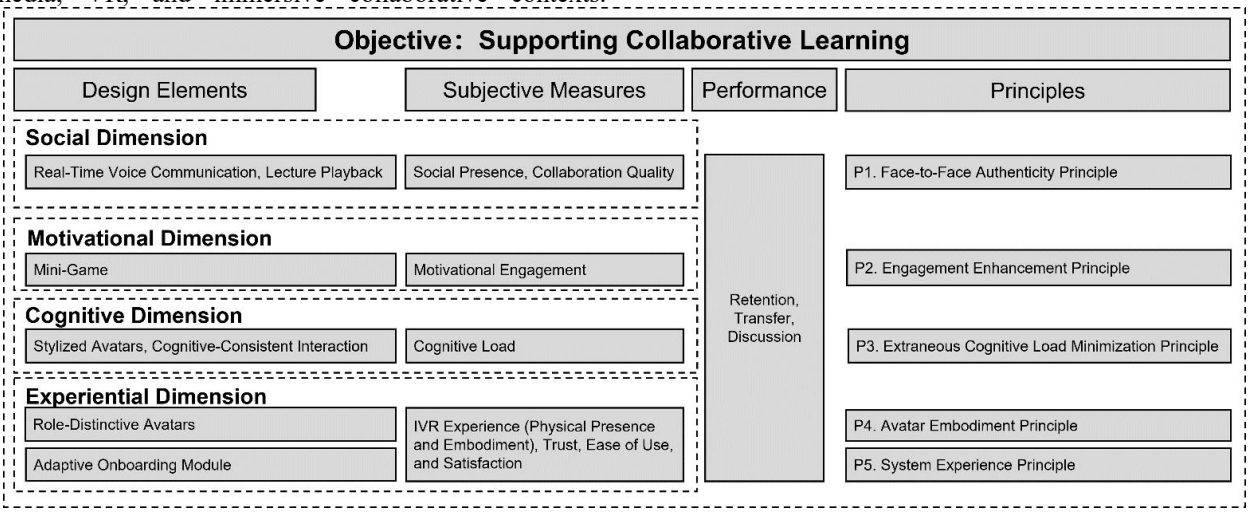


Fig. 1. Overview of MUIVE design elements, principles, and associated assessments (subjective and performance-based).

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

1) **Social Dimension**

Within the social dimension, two key design elements were implemented to support collaboration: real-time voice communication and lecture playback.

a) Real-Time Voice Communication

A central component of collaborative learning involves discussions among peers, through which learners exchange and build upon ideas related to instructional content. Most existing multi-user collaborative systems support communication via either voice or text [4], with text-based options including notes, messages, blogs, and similar tools. However, text input in MUIVE requires typing on a virtual keyboard using handheld controllers, which is a slow and cumbersome process that hinders the flow of communication. To overcome this limitation, we implemented Unity's Photon PUN 2 for multi-user networking and integrated Photon Voice to enable real-time voice communication (see Fig. 2(h)). This setup supports more natural interaction and ultimately enhances the quality of collaboration.

b) Lecture Playback

To support lecture video playback and review, we developed a multi-user video player using the AVPro plugin in Unity. This player allows multiple learners to collaboratively control playback using their hand controllers (see Fig. 2(g)). During group discussions, learners can interact with the video content in real time, such as pausing, rewinding, or replaying segments as needed. This functionality also aims to help reduce cognitive load by allowing learners to revisit content on demand.

To assess the social dimension, we focused on two constructs: social presence [33] and collaboration quality [34]. Social presence reflects learners' perceived sense of connection and interactivity during collaboration. When deeply engaged, learners often develop stronger social bonds, which in turn enhance learning outcomes [11]. Collaboration quality reflects learners' subjective evaluation of group work [35].

2) **Motivational Dimension**

To foster learner engagement from the outset, we introduced a warm-up game as a motivational design element. Research suggests that interactive and moderately challenging tasks can significantly boost learners' willingness to participate and collaborate [36]. In MUVEs, games are often used not only to help users familiarize themselves with system operations but also to increase peer connectedness. For example, Dindar et al. [36] emphasize that the sense of team cohesion fostered through gameplay can facilitate effective communication channels, promote information sharing, and enhance mutual understanding among team members. These outcomes, in turn, support more efficient collaboration by enabling timely exchange of ideas and feedback, while also reinforcing social presence and perceived collaboration quality.

In MUIVE, we designed a two-player relay mini-game (see Fig. 2(f)) in which learners work together to complete a simulated baton-passing task. The baton was implemented as a graspable object embedded with gravity physics to replicate real-world dynamics. This game was intended to create a playful and emotionally engaging shared experience, helping to build early-stage interpersonal rapport and collaborative cohesion. To assess this dimension, we examined learners' motivational engagement, focusing on their perceived enjoyment, involvement, and sense of engagement in the collaborative experience.

3) **Cognitive Dimension**

In immersive virtual learning environments, the design of system elements should carefully consider the cognitive load imposed on learners during collaboration [5]. Given that IVR systems typically integrate multiple streams of multimedia information, excessive visual or interactive stimuli may lead to redundancy and increase extraneous cognitive load [37]. To mitigate this, the collaborative environment in MUIVE was designed to present only essential information, minimize sensory distractions, and streamline interaction mechanisms, all while supporting core learning goals [28].

Two specific design strategies were employed to address cognitive load: stylized avatars and cognitively consistent interaction.

a) Stylized Avatars

To reduce visual complexity, learners were represented by minimalistic, cartoon-style avatars (see Fig. 2(c)), consisting only of the torso, head, and hands. Prior studies have suggested that highly realistic avatars may trigger the "uncanny valley" effect, potentially disrupting learning [38]. In contrast, simplified, stylized avatars have been shown to improve cognitive comfort and learning outcomes. Additionally, by omitting detailed limb modeling and motion simulation, the system reduces the computational and perceptual demands placed on users, thereby lowering extraneous cognitive load and enhancing the overall learning experience.

b) Cognitive-Consistent Interaction

To further reduce operational load, interaction design in MUIVE was aligned with users' everyday cognitive schemas [39]. Avatar movement was controlled via intuitive thumbstick input; video controls employed normal play and pause icons; speaking avatars were marked with a microphone symbol above their heads, while listeners were marked with a speaker icon (see Fig. 2(g)). Object manipulation followed real-world conventions: pressing the trigger to grasp and releasing it to drop. By embedding real-world interaction logic into the system, these features aim to minimize the learning curve and support faster mastery of system operations.

To assess this dimension, we focused on learners' cognitive load, including their mental load and effort.

4) **Experiential Dimension**

The experiential dimension focuses on enhancing learners' sense of immersion, embodiment, and overall system satisfaction within MUIVE. To support these experiential qualities, two key design elements were implemented: role-distinctive avatars and an adaptive onboarding module.

a) Role-Distinctive Avatars

To support different instructional contexts, MUIVE allows both student-led and teacher-guided collaborative learning. When instructor participation is required, teachers can designate their role during login by selecting the "Teacher"

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

label (see Fig. 2(a) and 2(b)). The system then assigns them a distinct avatar, visually differentiated from student avatars and clearly marked with a "Teacher" tag, to facilitate role recognition in the environment (see Fig. 2(g)). This design facilitates clear role identification, enabling instructors to monitor learners' progress, provide timely guidance, and intervene when necessary. By reinforcing role awareness and promoting structured interaction, this feature supports both engagement and the perceived quality of collaboration.

b) Adaptive Onboarding Module

To foster users' comfort and adaptability within the immersive system, an adaptive onboarding module was introduced at the entry stage. Adaptability here refers to the extent to which learners can gain control over the system environment and adjust their behaviors accordingly [40]. Upon entry, learners are provided with two assistive tools: a mirror (Fig. 2(c)) and a practice video player (Fig. 2(d)). The mirror, implemented via a virtual camera, allows learners to view their own movements reflected in their avatar in real time. This helps users align their physical gestures with the virtual representation, thereby enhancing embodiment and sensorimotor synchrony. The practice video player enables learners to explore basic playback controls and rehearse common actions, such as turning, moving, and clicking, before engaging in actual learning tasks. Together, these tools aim to improve system usability, reduce operational friction, and strengthen both physical presence and embodied experience.

To evaluate the experiential dimension, we assessed the following constructs: physical presence, which reflects learners' subjective sense of self-location and immersion within the virtual environment [41]; embodiment, defined as the perceived alignment and identification between one's physical body and the virtual avatar [42]; trust, referring to learners' perception of the system's stability and reliability, which is essential for sustained collaborative engagement [43]; ease of use, which captures the perceived learnability and operability of the system, that is, simple, intuitive interfaces can reduce cognitive load and increase users' focus and efficiency [44]; satisfaction, which reflects learners' overall appraisal of the system's functionality and support for collaborative learning. Higher satisfaction typically indicates a stronger match between the system design and learners' needs, leading to more positive engagement and acceptance [45].

B. Implementation

1) **The System**

MUIVE was developed using the Unity engine, together with Photon PUN 2, Photon Voice, and the AVPro plugin. The Photon PUN 2 module enables multi-user networking, while Photon Voice supports real-time voice communication to ensure fluid interaction and natural conversation during collaboration. The system also integrates the AVPro plugin to control lecture video playback, allowing learners to manage pacing and review content as needed.

Based on the conceptual framework, we developed the MUIVE system to support immersive multi-user collaborative learning. As shown in Fig. 2, users begin at the login interface, choosing either real-name or anonymous access (Fig. 2(a)). Real-name users enter their role and ID (Fig. 2(b)). They then complete an adaptation phase, which includes viewing their avatar via a mirror (Fig. 2(c)) and a brief lecture video practice to familiarize themselves with the controls (Fig. 2(d)). After adaptation, users enter the collaborative environment (Fig. 2(e)), which includes a game area to boost engagement (Fig. 2(f)), a lecture video module for content playback (Fig. 2(g)), and a discussion zone for peer interaction (Fig. 2(h)).

2) **Learning Materials**

The instructional content in this study consisted of basic concepts in artificial intelligence (AI). AI literacy has become increasingly essential for non-computer science majors, as understanding fundamental AI principles is now considered a critical component of general education in many disciplines [46]. However, the abstract nature of AI knowledge and the ethical issues it raises often present significant challenges for teaching and learning [47]. MUIVE delivers content covering machine learning, neural networks, deep learning, generative adversarial networks, Transformer models, and AI ethics. A blended instructional approach combining lecture-based delivery with collaborative learning was employed to promote conceptual understanding and encourage peer interaction. Lecture videos were created using a generative teaching video pipeline adapted from [48].

IV. User Study

A. Participants and Research Design

A total of 104 undergraduate and graduate students (52 males, 52 females; age range = 17-27, M = 20.78, SD = 2.61) were randomly recruited for the study (see Table I). Participants were recruited through campus-wide public recruitment using both online and offline approaches. Study advertisements were posted in campus group chats and online forums, and research assistants also conducted in-person recruitment in public campus spaces such as the basketball court and cafeteria. To be eligible, participants were required to be from majors unrelated to computer science and to report no self-reported susceptibility to VR sickness. All participants provided written informed consent and received 35 RMB upon completion of the study. No sample attrition occurred during the experiment. A between-subjects design was adopted, with the learning approach (MUIVE-based online synchronous learning and video-based online synchronous learning) as the independent variable. Participants were randomly assigned to either the experimental group, which used the MUIVE system, or the control group, which used a conventional online collaborative learning platform, Tencent Meeting. As a widely adopted platform for student collaboration, Tencent Meeting supports common collaborative learning activities such as voice discussion, text-based communication, shared viewing of slides and videos.

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

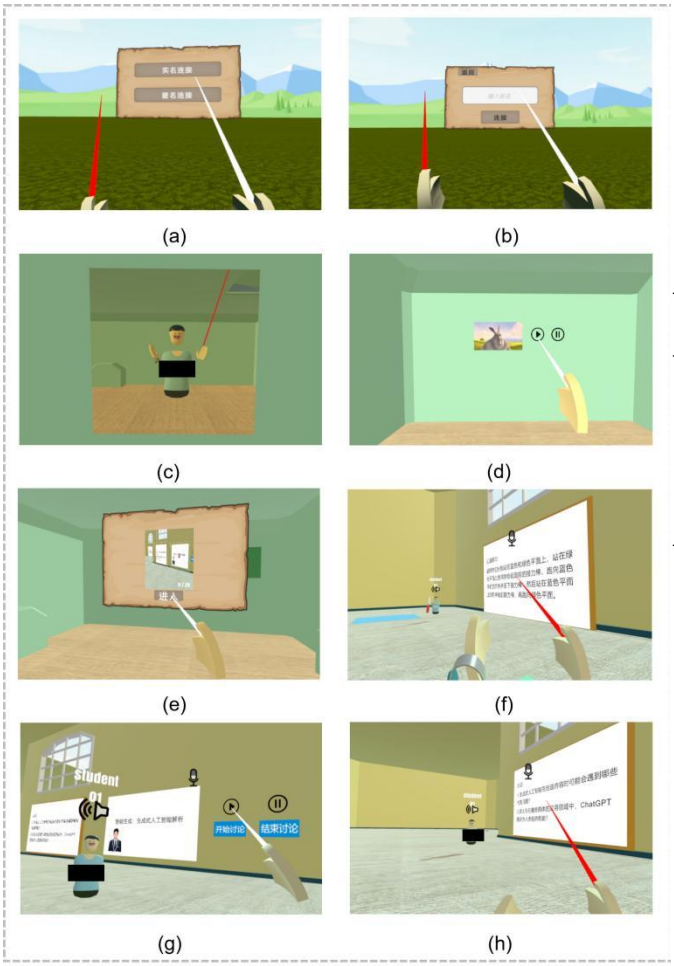


Fig. 2. Overview of the MUIVE system interface and user flow. (a) Login interface with real-name and anonymous access options. (b) Real-name login requires role and ID input. (c) Avatar adaptation through a mirror view. (d) Lecture video practice for interface familiarization. (e) Entry into the multi-user collaborative environment. (f) Game area for engagement. (g) Lecture video module for content playback. (h) Discussion zone for peer interaction.

TABLE I
PARTICIPANT ALLOCATION BY GROUP

	Experimental group	Control group
Male	26	26
Female	26	26
Male dyad	12	12
Female dyad	12	12
Mixed-gender dyad	2	2

To ensure peer learning conditions, participants worked in dyads (two-person learning groups). Group assignment was randomized while balancing gender and matching partner familiarity levels. The full group distribution is shown in Table I. A pre-test assessed prior knowledge of AI. A dyad-level independent-samples t-test showed no significant difference in total pre-test scores between groups ($t = -1.376$, $p = .175$, Cohen's $d = -0.382$). Partner familiarity was measured using a 5-point Likert scale. The ratings of the two members within each dyad were averaged to obtain a dyad-level familiarity score. A Mann–Whitney U test showed no significant difference in dyad-level familiarity between the two groups ($U = 325.000$, $Z = -0.246$, $p = .806$, $r = .034$) (see Table II).

TABLE II
DESCRIPTIVE STATISTICS FOR GROUPS

	Experimental group		Control group			
	M	SD	M	SD	Statistic	p
Age (Individual)	21.00	2.72	20.56	2.52	Z = -0.519s	.604
Gender (Individual)	1.50	0.51	1.50	0.51	$\chi^2 = 0.000$	1.000
Familiarity level (Dyad)	2.96	1.67	3.15	1.57	U = 325.000, Z = -0.246	.806

Note: Individual = individual-level variable; Dyad = dyad-level variable. Gender coding: 1 = Male; 2 = Female.

To ensure consistency in learning content, all participants watched the same instructional video. The experimental group viewed the video and discussed it collaboratively within the MUIVE platform. The control group watched the same video via screen sharing on a standard conferencing platform and discussed the same two questions.

B. Procedure

Fig. 3 presents the experimental procedure. All sessions took place in a quiet room. Participants first received an overview of the study, signed informed consent, and completed a brief demographic questionnaire (e.g., age, gender, major) along with a pre-test. The MUIVE system was operated on a laptop connected to a VR headset. Participants in the experimental group began in a Home scene, where they practiced basic interactions such as video playback. They were then guided to the Learning scene, which included a Game area and a Learning area. In the Game area, participants completed a cooperative relay mini-game to familiarize themselves with system interactions. They then moved to the Learning area to watch an instructional video and discuss two assigned questions. Control group participants used an online learning platform to view the same video and discuss the same questions collaboratively. After the learning session, all participants completed two open-ended questions, the post-test, a subjective experience questionnaire, and a brief interview. Each session lasted approximately 45 minutes.

C. Experimental Setup

The MUIVE was operated using an HTC VIVE Pro, which includes a head-mounted display (HMD), controllers, and tracking base stations. The HMD offers a combined resolution of 2880 × 1600, a 90 Hz refresh rate, 64 mm interpupillary distance, and built-in high-resolution headphones. SteamVR Tracking was used for spatial tracking, and interaction was handled via VIVE controllers. The Unity application was connected to the device through the SteamVR plugin, enabling real-time scene control and interaction.

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

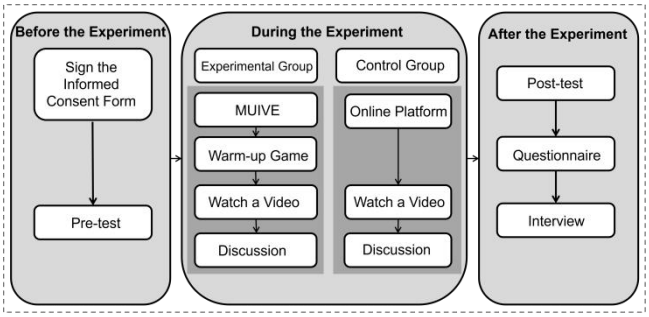


Fig. 3. The experimental procedure

D. Learning Performance

To evaluate learning outcomes, two AI experts developed a set of test items and discussion prompts based on the instructional content. These served as pre- and post-test instruments. The test items measured knowledge retention (recall of factual content) and knowledge transfer (application in new contexts), while the discussion prompts assessed collaborative learning performance.

1) **Pre-test**

The pre-test assessed participants' baseline familiarity with AI and their initial ability to transfer knowledge. It included two sections: a retention test (maximum score = 9), which evaluated factual recall (e.g., "What is the definition of machine learning?"), and a transfer test (maximum score = 9), which measured the application of knowledge in new contexts (e.g., "Which of the following tasks is most likely to involve unsupervised learning?"). Each item was scored as 1 for a correct response and 0 for an incorrect one. Lower overall scores reflected weaker prior knowledge and transfer ability.

2) **Post-test**

The post-test consisted of two parts: open-ended questions and a knowledge post-test. In the open-ended section, participants were asked to discuss in pairs and collaboratively respond to two prompts: (1) What ethical issues might arise when generative AI is used to create content? and (2) In which specific application areas can ChatGPT support human work? The knowledge post-test contained the same retention and transfer components as the pre-test and was used to assess both knowledge retention and knowledge transfer. The pre-test and knowledge post-test were identical, with the same score ranges for retention (0-9) and transfer (0-9), allowing direct comparison of performance changes over time. The test demonstrated acceptable internal consistency, with a Cronbach's α of 0.65, exceeding the commonly accepted threshold of 0.60 [49].

3) **Gain scores**

Gain scores were calculated to assess changes in learning performance from the pre-test to the post-test. Retention and transfer gain scores were first calculated at the individual level by subtracting each participant's corresponding pre-test score from the post-test score (gain score = post-test score − pre-test score). The gain scores of the two members within each fixed dyad were then averaged to obtain dyad-level retention and transfer gain scores for inferential analysis. The overall dyad-level gain score was calculated as the sum of the retention and transfer gain scores. Higher gain scores indicated greater improvement in learning performance.

E. Learning Experience

Subjective learning experiences were assessed using validated self-report psychometric scales across four dimensions: social, motivational, cognitive, and experiential. These dimensions encompassed nine measures: social presence, collaboration quality, motivation, cognitive load, physical presence, embodiment, trust, ease of use, and satisfaction. The questionnaire included a total of 72 items, each rated on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree). Higher scores reflected greater agreement with the corresponding statements. All measures were adapted from established instruments reported in prior research, with the relevant references summarized in Table III. To fit the context of the present study, only minor wording adaptations were made. For example, an original item such as "*I enjoyed this type of learning system very much*" was adapted for the present context as "*I enjoyed this multi-user virtual reality learning system very much.*" The full list of subjective measures is reported in Appendix A (Table A1). The internal consistency of the adopted scales was acceptable to good, with all Cronbach's α values exceeding .70.

TABLE III
SUBJECTIVE MEASURES

Dimension	Measure	Numbers of items	Type	Reliability (Alpha)	Reference
Social	Social presence	5	7-point Likert	.875	[41]
	Collaboration quality	20		.894	[35], [50]
Motivation	Motivation	8		.901	[45]
Cognitive	Cognitive load	8		.896	[44]
Experiential	IVR experience (Physical Presence and Embodiment)	12		.985	[42], [41]
	Trust	5		.798	[43]
	Ease of use	7		.934	[44]
	Satisfaction	7		.825	[45]

F. Brief Interview

The purpose of the interview data collection and analysis was to help explain the observed effects and provide insights for the future design of MUIVE. After completing the subjective questionnaires, participants were asked to provide suggestions and comments about the collaborative learning platform they had just used. This was done by responding to

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

an open-ended prompt: "Please provide any suggestions or comments you have about this collaborative learning platform."

Written responses were analyzed using open coding. Two independent coders initially reviewed all participant responses and independently generated preliminary codes. Following this phase, the coders met to compare and reconcile coding differences through discussion, thereby reaching a consensus and producing a finalized coding scheme. This iterative process ensured intercoder reliability and comprehensiveness of thematic coverage. Through this analysis, eight themes were identified: collaborative experience, perception of the learning environment, learning outcome feedback, preference for learning resources, cognitive load issues, interaction experience, technical issues, and other feedback. These themes served as the analytical framework for analyzing all participant responses.

To assess the reliability of the coding process, Cohen's Kappa [51] was used to calculate inter-rater agreement. The analysis showed a Kappa value of 0.815, indicating a high level of agreement between coders [52]. The result was statistically significant ($p < .001$), confirming strong consistency. After resolving discrepancies through discussion, full consensus was achieved on the final coding scheme.

G. Data Analysis

Data analysis was conducted using SPSS. Because participants completed the collaborative learning activities in fixed two-person dyads and the experimental condition was implemented at the dyad level, independence of observations within dyads could not be assumed a priori. To empirically characterize the degree of within-dyad interdependence, single-measure intraclass correlation coefficients (ICCs) were estimated using a one-way random-effects model prior to aggregation. The ICC values for the primary outcomes are reported in Tables IV and V and indicate that the degree of within-dyad similarity varied across outcomes. Given the dyadic structure of the experimental design and collaborative learning activities, individual quantitative outcomes were averaged across the two members of each dyad for between-condition inferential analyses [53]. This procedure yielded 52 dyad-level observations, including 26 dyads in the MUIVE condition and 26 dyads in the control condition.

Normality was assessed via statistical tests and Q-Q plots to determine appropriate methods. For normally distributed data, independent-samples t-tests were used, with effect sizes reported as Cohen's d (small = 0.20, medium = 0.50, large = 0.80 [54]). For non-normal data, Mann-Whitney U tests were applied, with effect sizes reported as r (small = 0.10, medium = 0.30, large = 0.50 [55]). As a sensitivity analysis, dyad-level analyses of covariance (ANCOVA) were additionally conducted to examine whether group differences in post-test performance remained after controlling for corresponding pre-test scores. Pearson correlations between pre-test and post-test scores were calculated at the dyad level to examine the relationship between baseline and post-intervention performance.

V. Results

A. Performance

Descriptive statistics and effects of performance are reported in Table IV. The difference in knowledge retention gain scores between the two groups was not statistically significant ($t = 1.697$, $p = .096$, Cohen's $d = 0.471$).

TABLE IV
STATISTICS OF LEARNING PERFORMANCE

Test type	Dimension	Experimental M (SD)	Control M (SD)	t	p	Cohen's d	ICC
Pre-test	Retention	0.58 (0.90)	0.89 (0.78)	-1.317	.194	-0.365	.102
	Transfer	0.73 (0.78)	1.02 (1.13)	-1.074	.288	-0.298	.445
	Overall	1.31 (1.47)	1.90 (1.65)	-1.376	.175	-0.382	.389
Post-test	Retention	7.12 (0.95)	6.89 (1.13)	0.798	.428	0.221	.001
	Transfer	5.54 (1.48)	5.12 (1.53)	1.014	.315	0.281	.371
	Overall	12.65 (2.28)	12.00 (2.23)	1.044	.301	0.290	.242
Gain scores	Retention	6.54 (1.06)	6.00 (1.22)	1.697	.096	0.471	.115
	Transfer	4.81 (1.50)	4.10 (1.36)	1.791	.079	0.497	.391
	Overall	11.35 (2.28)	10.10 (2.15)	2.030	.048	0.563	.351

With respect to knowledge transfer gain scores, the difference was not statistically significant between the two groups ($t = 1.791$, $p = .079$, Cohen's $d = 0.497$).

As shown in Fig. 4, an independent-samples t-test revealed a significant difference between the experimental and control groups in the overall pre-post test score gains ($t = 2.030$, $p = .048$, Cohen's $d = 0.563$), indicating a moderate effect size. To further examine whether the observed performance differences remained after controlling for baseline performance, dyad-level ANCOVAs were conducted as a sensitivity analysis. The interaction between condition and corresponding pre-test scores was not significant for retention, transfer, or overall performance, indicating that the homogeneity of regression slopes assumption was satisfied. After controlling for pre-test scores, the effect of condition was not significant for retention ($F(1,49) = 1.389$, $p = .244$, $\eta p^2 = .028$), transfer ($F(1,49) = 2.345$, $p = .132$, $\eta p^2 = .046$), or overall performance ($F(1,49) = 2.640$, $p = .111$, $\eta p^2 = .051$).

In addition, Pearson correlation analyses were conducted at the dyad level to examine the relationship between pre-test and post-test performance. The results showed a positive association between baseline and post-intervention performance. Specifically, retention scores showed a positive but non-significant correlation ($r = .250$, $p = .074$), whereas

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

transfer scores (r = .365, p = .008) and overall performance scores (r = .331, p = .016) were significantly correlated.

B. Subjective Measures

Descriptive statistics and effects of subjective measures are reported in Table V. IVR experience was measured only in the MUIVE group; therefore, no between-group comparison was conducted.

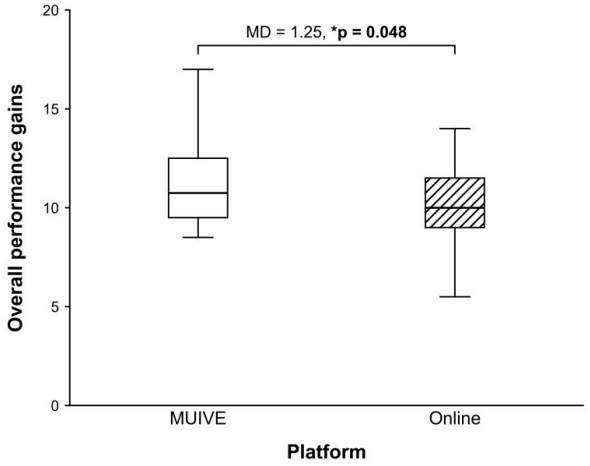


Fig. 4. Differences in the overall performance gains between the two groups

1) **Social Dimension**

In terms of social dimension, two subdimensions were assessed: social presence and collaboration quality. As illustrated in Fig. 5, independent-samples t-test results indicated a significant difference in social presence between the two groups (t = 2.468, p = .017, Cohen's d = 0.684). However, no significant difference was observed in collaboration quality (t = 1.911, p = .062, Cohen's d = 0.530). When these two subdimensions were combined into an overall index of the social dimension, the experimental group still outperformed the control group significantly (t = 2.305, p = .025, Cohen's d = 0.639), indicating that the immersive environment fostered a more socially engaging learning experience.

2) **Motivational Dimension**

As shown in Fig. 6, an independent-samples t-test revealed a significant difference in learning motivation between the two groups (t = 4.700, p < .001, Cohen's d = 1.303).

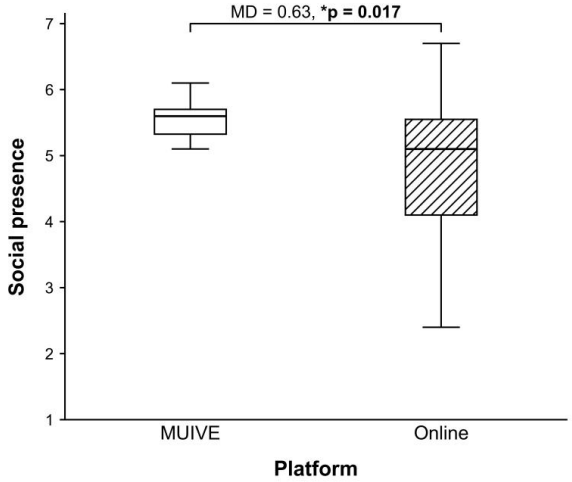


Fig. 5. Differences in social presence between the two groups

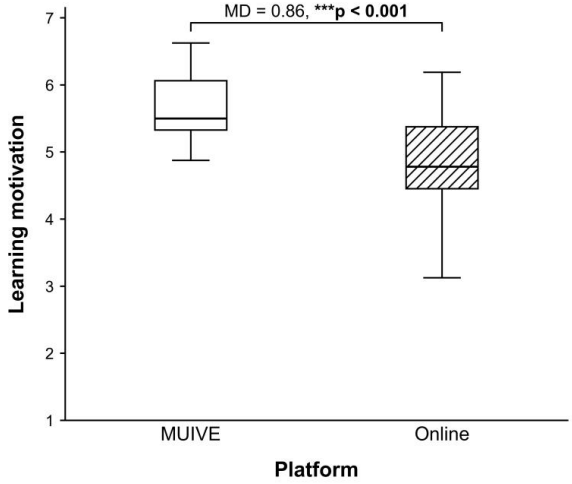


Fig. 6. Differences in learning motivation between the two groups

TABLE V
STATISTICS OF SUBJECTIVE MEASURES

Dimension	Measure	Experimental Group M(SD)	Control Group M(SD)	t	p	Cohen's d	ICC
Social	Social Presence	5.44 (0.77)	4.81 (1.05)	2.468	.017	0.684	.166
	Collaboration Quality	5.50 (0.42)	5.25 (0.55)	1.911	.062	0.530	-.053
Motivation	Learning Motivation	5.64 (0.47)	4.78 (0.81)	4.700	<.001	1.303	.092
Cognitive	Cognitive Load	5.20 (0.79)	5.03 (0.76)	0.773	.443	0.214	-.119
	IVR Experience	5.03 (0.60)					-.255
Experiential	Trust	5.35 (0.50)	4.86 (0.83)	2.627	.011	0.729	-.002
	Ease of Use	5.95 (0.44)	5.33 (0.62)	4.168	<.001	1.156	-.106
	Satisfaction	5.69 (0.53)	5.27 (0.57)	2.750	.008	0.763	-.133

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

3) **Cognitive Dimension**

As shown in Fig. 7, an independent-samples t-test revealed no significant difference in cognitive load between the two groups (t = 0.773, p = .443, Cohen's d = 0.214), indicating that the MUIVE system did not increase students' cognitive load.

4) **Experiential Dimension**

The experiential dimension was assessed through participants' IVR experience, perceived trust, ease of use, and satisfaction with the system. The IVR experience was assessed through two subdimensions: physical presence (M = 4.98, SD = 0.74) and embodiment (M = 5.05, SD = 0.67). The mean scores for both subdimensions exceeded the scale midpoint, indicating that learners experienced a relatively strong sense of physical presence and embodiment in MUIVE. The overall IVR experience score was also above the midpoint (M = 5.03, SD = 0.60), suggesting that participants generally reported a positive immersive experience.

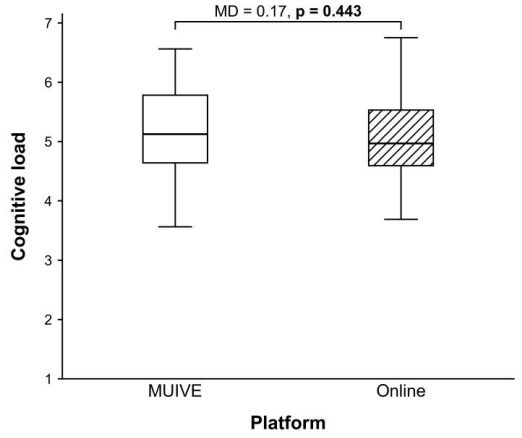


Fig. 7. Differences in cognitive load between the two groups

Independent-samples t-tests revealed significant group differences in trust, ease of use, and satisfaction (see Fig. 8). Trust in the system was significantly higher in the experimental group compared with the control group (t = 2.627, p = .011, Cohen's d = 0.729). Ease of use also showed a significant difference between groups (t = 4.168, p < .001, Cohen's d = 1.156). Satisfaction, as a core indicator of experiential quality, also differed significantly between groups (t = 2.750, p = .008, Cohen's d = 0.763).

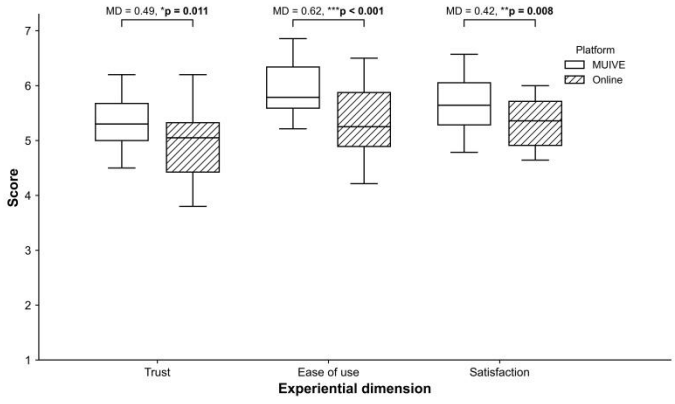


Fig. 8. Differences in experiential dimension between the two groups

5) **Open-ended Responses**

After the IVR learning and discussion sessions, participants were instructed to respond in writing to two open-ended questions using pen and paper. The responses were independently evaluated by two domain experts in AI. To assess the reliability of the scoring, inter-rater consistency was examined. The internal consistency analysis yielded a high Cronbach's α of .912 for the two items, indicating excellent reliability across responses [56]. Additionally, inter-rater agreement was assessed using a two-way random-effects model to compute the intraclass correlation coefficient (ICC). The results indicated strong consistency, with a single-measure ICC of .834 and an average-measure ICC of .910 [57]. An independent-samples t-test revealed no significant difference in performance between the two groups (t = -0.436, p = .665, Cohen's d = -0.121).

C. Impact of Peer Familiarity on Performance

To examine whether the effect of platform on performance varied by peer familiarity, the original 5-point partner-familiarity ratings were recoded into a binary variable distinguishing familiar and unfamiliar dyads, and regression analyses were conducted at the dyad level. For knowledge retention, neither familiarity (B = -0.406, p = .381) nor its interaction with platform type (B = 0.491, p = .462) reached statistical significance. For knowledge transfer, the effect of familiarity within the MUIVE condition was not statistically significant (B = -0.806, p = .148). Nevertheless, dyads composed of unfamiliar peers showed higher transfer gains (M = 5.27, SD = 1.63) than those composed of familiar peers (M = 4.47, SD = 1.36). Conversely, in the conventional online platform, familiar dyads showed descriptively higher transfer gains (M = 4.47, SD = 1.19) than unfamiliar dyads (M = 3.39, SD = 1.43). More importantly, the interaction between familiarity and platform was significant and positive (B = 1.888, p = .021). Among unfamiliar dyads, the conventional online platform yielded significantly lower transfer gains than MUIVE (B = -1.884, p = .004), whereas familiar dyads demonstrated nearly identical transfer gains across the two platforms. These findings indicate that the effect of platform on transfer performance depended on peer familiarity, with the advantage of MUIVE primarily concentrated among unfamiliar dyads.

Regarding overall learning performance, the effect of familiarity within the MUIVE condition was not statistically significant (B = −1.212, p = .169), although unfamiliar dyads showed higher overall gains (M = 12.05, SD = 2.29) than familiar dyads (M = 10.83, SD = 2.22). In the conventional online platform, familiar dyads showed descriptively higher overall gains (M = 10.50, SD = 1.77) than unfamiliar dyads (M = 9.33, SD = 2.69). Among unfamiliar dyads, overall gains were significantly lower in the conventional online platform than in MUIVE (B = −2.712, p = .008). However, the interaction between platform and peer familiarity did not reach statistical significance (B = 2.379, p = .063). Thus, although overall performance gains showed a crossover pattern similar to that observed for transfer performance, there was

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

insufficient evidence that the effect of platform on overall performance varied significantly according to peer familiarity.

D. Interview Data Analysis

To further interpret the quantitative findings, representative interview comments were organized into four dimensions: social, motivational, cognitive, and experiential. The original Chinese comments and their English translations are reported in Table A2 in Appendix A. Overall, the interview data suggest that MUIVE was perceived as more socially engaging, motivating, and experientially rich, whereas the learning burden reported by participants was more closely related to the instructional content and its presentation than to the collaborative system itself.

In the social dimension, comments (P26, P28, and P42) suggested that MUIVE reduced awkwardness and made collaboration feel closer to real-life interaction. In addition, one participant (P2) noted that peer interaction in MUIVE helped support understanding of difficult concepts. By contrast, participants in the control group more often described Tencent Meeting as lacking embodiment, social presence, and a genuine sense of interaction (P85 and P87). These responses indicate that MUIVE was perceived as more supportive of social connectedness and collaboration initiation.

In the motivational dimension, participants in the experimental group described MUIVE as interesting, enjoyable, and engaging (P10, P35, and P47). These comments were related to system novelty, interaction, and embodiment. In contrast, participants in the control group (P97 and P99) more often referred to the learning as distracting or dull, and one participant (P99) explicitly expected VR to be more novel and conducive to concentration.

In the cognitive dimension, the burden reported by participants did not appear to stem primarily from the collaborative system itself. Instead, comments from both groups pointed to the abstractness of the content, the rigid style of explanation, and learners' unfamiliarity with the topic (P3, P16, and P79). This pattern suggests that the perceived cognitive load was shaped more by the materials and their presentation than by the collaborative environment alone.

In the experiential dimension, participants in the experimental group generally valued the system's ease of use, novelty, and immersive quality (P2, P4, and P24). Participant P24 also noted that the adaptive module helped the learner become familiar with the system quickly, making subsequent tasks easier. Participant P18 reported limitations related to headset use, such as dizziness and insufficient visual clarity. These comments suggest that MUIVE was experienced as novel, usable, and immersive, while also revealing some device-related constraints in the overall user experience.

VI. Discussion

A. Empirical Contributions

The gain-score analysis showed a significant advantage for MUIVE in overall performance gains, although no significant group differences were observed in retention gains, transfer gains, or open-ended performance. However, the baseline-adjusted ANCOVA did not reveal significant condition effects on post-test retention, transfer, or overall performance. Taken together, these findings suggest a positive but not fully consistent performance pattern for MUIVE. This pattern may be related to MUIVE's advantages in social, motivational, and experiential aspects, including higher social presence, motivation, trust, ease of use, and satisfaction. Importantly, these benefits were achieved without imposing additional cognitive load. Beyond the comparison with Tencent Meeting, this finding is also meaningful in relation to other virtual learning environments. Compared with conventional online platforms, immersive multi-user virtual environments provide stronger embodied and spatial affordances [29]. Prior research also suggests that, compared with desktop VR and other less immersive environments, immersive collaborative environments can better support learning when key affordances such as social presence, spatial information, embodied interaction, and collaboration are incorporated into the learning design [11], [27].

At the social level, MUIVE fosters a significantly higher sense of social presence than the conventional online instructional platform. Unlike traditional platforms that rely on voice or video calls and lack the rich contextual and non-verbal cues essential for authentic interpersonal connection [58], MUIVE provides a face-to-face-like collaborative experience that amplifies the feeling of "being with others" [59]. This interpretation is in line with recent theoretical work on immersive collaborative learning, which suggests that immersive environments differ from traditional computer-mediated collaboration by affording stronger social presence, physical presence, body ownership, and agency, all of which can shape collaborative interaction and, ultimately, learning outcomes [11]. This advantage is partly enabled by the immersive setting, which conveys real-time embodied cues such as peer location, gaze direction, and hand gestures that are crucial components of natural human interaction. As Sterna and Zibrek emphasized [60], such detailed and synchronous nonverbal signals are critical for establishing social presence. In MUIVE, learners perceive and respond to these signals, which reinforce the sense of connection among peers [61]. Qualitative evidence from the interviews also supports this view: several participants described the virtual setting as reducing awkwardness and making collaboration feel closer to real-life interaction, particularly when working with unfamiliar peers (e.g., P26, P28, and P42). In addition, one participant noted that although some concepts were difficult to understand, peer interaction within MUIVE helped support comprehension, further highlighting the social value of the collaborative environment (P2).

In terms of motivation, the embedded warm-up game helped build dyadic rapport and encouraged learners' willingness to participate. For instance, Participant P10 reported, "The warm-up game made me more eager to participate in the learning activity and brought me closer to my partner." In addition to helping learners quickly master system controls, the game's intrinsic playfulness served as an effective motivational trigger, which in turn promoted more active and higher-quality peer interaction during the main collaborative task. Interview findings also suggest that participants perceived MUIVE as enjoyable, interesting, and engaging (e.g., P35 and P47).

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

Regarding the IVR experience, learners reported strong feelings of embodiment and physical presence within MUIVE. For example, Participant P4 described the system as highly immersive and the learning experience as face-to-face-like and satisfying. According to embodied cognition theory, cognitive processing is deeply rooted in bodily actions and sensorimotor experiences [62]. Interacting physically with the virtual space enabled learners to internalize conceptual relationships and task structures more deeply, thereby supporting cognitive construction. Upon entry into MUIVE, learners engage with an adaptive orientation module, featuring a virtual mirror that synchronizes facial and body movements in real time. This enhanced sense of embodiment prepares learners before instruction begins. Additionally, physical presence, the sense of "being there" in the virtual space [63], was strengthened through design elements such as embodied peer interaction (e.g., hand gestures, body orientation). Learners also customized avatar appearance and name tags based on assigned roles (e.g., teacher or student), further deepening role identification and immersive engagement.

Trust refers to learners' perception of a system's reliability and stability [64]. In MUIVE, real-time visual indicators, such as a microphone icon in front of the speaker and a speaker icon above the listener's avatar, provide immediate and accurate feedback, which enhances trust. For time-sensitive tasks, system responsiveness is critical. To reduce latency, Photon PUN 2 servers hosted domestically were selected, ensuring stable real-time interaction. Additionally, established icons (e.g., triangle for play, two bars for pause) aligned with learners' habitual expectations, further strengthening their confidence in the platform. Ease of use captures how easy the technology is perceived to be by learners [65]. MUIVE's adaptive onboarding module enabled learners to quickly grasp core functions. Interview comments also supported the positive experiential evaluations of MUIVE. For example, Participant P2 described the system as simple and easy to understand, with its functions being mastered in a short time, while Participant P24 emphasized the novelty of the system and noted that the adaptive module helped the learner become familiar with the system quickly, which made subsequent tasks easier. High learner satisfaction also stemmed from the system's instructional design, which seamlessly integrates lecture delivery with peer discussion. This design reflects real-world classroom practice, beginning with knowledge input and followed by collaborative construction.

From a cognitive perspective, MUIVE did not impose additional cognitive load relative to the conventional online platform. This finding is noteworthy because MUIVE involved more spatial and interactive operations than Tencent Meeting, including avatar selection, mirror-based orientation, three-dimensional navigation, and the relay mini-game. While IVR environments are known for rich multimodal content, which can increase cognitive demands, MUIVE did not abandon the core design principles of the cognitive theory of multimedia learning [66], but rather extended them through a minimalist interface and adaptive strategies to further reduce extraneous cognitive load. For instance, avatars were rendered with only essential components (head, torso, hands) to reduce complexity, and task instructions appeared on clean, distraction-free panels. The adaptive onboarding module also supported system navigation, reducing learners' extraneous load. Moreover, most of these additional operations occurred before the formal learning task and served as guided practice, while the simplified interface and familiar controls helped offset their operational demands. Additionally, the interview data suggest that the cognitive burden reported by learners was related less to the collaborative system itself than to the abstractness of the content and the way it was presented (e.g., P3, P16, and P79).

Importantly, MUIVE also reduced collaboration barriers caused by unfamiliar peer groupings. In the MUIVE condition, dyads composed of unfamiliar peers showed higher overall and transfer gains than those composed of familiar peers, although these within-platform differences were not statistically significant. In contrast, the control group showed the opposite descriptive pattern, with familiar dyads achieving higher gains than unfamiliar dyads. The significant interaction between platform and peer familiarity for transfer performance further indicated that the relative advantage of MUIVE was primarily evident among unfamiliar dyads. A similar crossover pattern was observed for overall performance, although the corresponding interaction did not reach statistical significance. The significant interaction for transfer performance suggests that MUIVE may be particularly beneficial when learners collaborate with unfamiliar peers. According to social penetration theory, interpersonal trust and communication depth develop gradually through information exchange [67]. In the early stages of collaboration, learners often withhold opinions, minimize interaction, and avoid self-disclosure, thereby impeding effective collaboration. The descriptive pattern observed in the control group was consistent with this tendency. However, MUIVE altered this pattern, particularly for transfer performance. A plausible explanation concerns the psychological safety experienced by unfamiliar partners. Psychological safety can be understood as a state of reduced interpersonal risk that enables individuals to participate more openly in interdependent activities [68]. Avatar-mediated interaction may be particularly beneficial for unfamiliar dyads because it reduces direct exposure of learners' physical appearance and may alleviate concerns about being negatively judged. Recent research in synchronous online learning has shown that avatar-based interaction can reduce students' interaction concerns, negative emotions, and social anxiety compared with webcam-based communication [69]. Such protection may be especially important for unfamiliar partners, who have not yet established interpersonal trust or stable communication routines. Familiar dyads, by contrast, may already rely on their existing relationships and therefore receive less additional benefit from avatar mediation.

Whereas conventional collaboration may inhibit participation because learners fear being judged negatively [70], MUIVE's immersive and embodied features may reduce evaluative anxiety, foster more open self-representation, and facilitate faster rapport-building among unfamiliar peers. As a result, MUIVE may provide particular support for collaborative engagement when learners have no established interpersonal relationship.

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

B. MUIVE Design Principles and Practical Implications

Drawing on empirical evidence from the current study, we propose five core design principles for the development of MUIVEs (see Table A3 in Appendix A), each offering practical insights into how future systems can better support collaborative learning. These principles are not intended to replace existing multimedia learning or educational VR design principles. Rather, they are grounded in key design concerns identified from prior studies, including social, motivational, cognitive, experiential, technological, and satisfaction-related aspects. In this sense, the contribution of the present study lies not in proposing principles detached from prior work, but in articulating a more focused principle set for MUIVEs based on both prior research and the empirical findings of this study.

Table A3 in Appendix A illustrates how the five principles relate to prior studies. The point here is not to suggest that earlier multimedia learning or educational VR principles are inadequate, but to show that many relevant design concerns have remained distributed across different strands of research. The five principles proposed here bring these concerns together more explicitly within the specific context of immersive multi-user collaborative learning. The Face-to-Face Authenticity Principle emphasizes simulating authentic co-located collaboration through multimodal interaction so that virtual collaboration more closely resembles face-to-face communication. Prior studies have addressed participation, social interaction [29], social presence, and physical presence [11]. The added value of this principle lies in bringing voice, gesture, peer visibility, and spatial positioning together under a unified design objective of authentic collaboration, thereby enhancing learners' sense of social presence and making virtual collaboration feel more natural and engaging. The Engagement Enhancement Principle advocates embedding playful or gamified elements into the system to boost learner motivation and participation. A warm-up game not only increased users' willingness to engage but also served as an onboarding mechanism, fostering a stronger sense of connection between unfamiliar partners before formal learning tasks. Existing work has discussed gamification [27], shared goals, scaffolds [29], and collaboration mechanisms [71], yet has paid limited attention to the initial phase of collaboration. This principle therefore extends engagement design to collaborative entry and links it to peer familiarization. The Extraneous Cognitive Load Minimization Principle highlights reducing unnecessary cognitive demands through interface simplification. This includes minimizing visual complexity, such as using simple avatar forms, and applying familiar interactive conventions, such as intuitive icons and controls, to ease user interaction and reduce extraneous processing. Prior work has examined cognitive load optimization [27], scaffolding [71], and overload and distraction control [72]. The present study further specifies how such concerns can be translated into concrete multi-user IVR features, such as stylized avatars, minimalist collaborative scenes, and cognitively consistent controls. The Avatar Embodiment Principle focuses on strengthening learners' sense of presence and role identification. By allowing avatar customization and incorporating features such as a real-time virtual mirror, learners can develop a stronger embodied connection with their avatar, reinforcing immersion and identity within the learning environment. Prior studies have discussed embodiment [71], body ownership, and agency [11]. This study advances these concerns to the level of system design by linking avatars to self-location, role recognition, and collaborative participation. Finally, the System Experience Principle highlights the value of an adaptive onboarding module. Such a module helps learners quickly become comfortable with the system's interface and functions, thereby reducing initial barriers to participation and allowing more cognitive resources to be allocated to learning tasks. Prior work has addressed scaffolding support [71] and reusability as quality criteria [72]. This principle reframes them as a continuous condition shaping collaborative participation rather than as isolated usability concerns.

These comparisons help clarify why the present study proposes a new set of principles. Existing research has provided many valuable design insights, but these insights have not yet been explicitly organized around the specific demands of immersive multi-user collaborative learning. The five principles proposed here therefore offer a more focused and context-specific framework for guiding the design of MUIVEs.

VII. Limitations and Future Directions

While this study designed and examined the effectiveness of MUIVE in supporting collaborative learning, several limitations remain. First, the proposed multidimensional framework was evaluated through MUIVE as a single system instantiation. Therefore, the findings primarily provide evidence for the effectiveness of this specific implementation rather than for the broader validation of the framework across different systems or contexts. Future research should further examine the transferability and broader applicability of the proposed framework across different platforms, contexts, and design instantiations. Second, some higher-level design features, such as visual learning cues, real-time feedback, and assessment mechanisms, were not included in the current version. These elements are crucial for scaffolding and regulating the learning process, and future work should extend the current design principles to incorporate them. Third, the present study was a short-term pretest–posttest study without longitudinal follow-up. Future research should explore the long-term effects of MUIVE on collaborative learning outcomes through longitudinal studies. Fourth, the number of dyads was relatively limited, which may have constrained the statistical power to detect moderate effects at the dyad level. In addition, the relatively low pre-test scores suggest the possibility of a floor effect, which may have limited the sensitivity of the pre-test to individual differences in baseline knowledge. Therefore, non-significant findings, particularly those related to learning performance, should be interpreted cautiously. Moreover, aggregating individual responses at the dyad level may obscure within-dyad variability in constructs such as cognitive load and learning motivation. Future research should employ larger samples and analytical approaches that better account for both dyad-level dependence and individual-level differences. Finally, the effectiveness of

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

MUIVE may vary across learners due to individual differences. Future studies should further investigate how traits such as spatial ability influence learning with MUIVE, providing insights for the development of adaptive systems that personalize the collaborative learning experience.

Ethics approval statement: This anonymized version omits identifying details for peer review purposes. This study received ethics approval [No. 2024-161] from the authors' institution (details omitted for peer review; available in the Title Page). The study complies with the WMA Declaration of Helsinki regarding participant recruitment, data collection, data use, data confidentiality, and protection of participants' rights. All participants were informed of the voluntary nature of their participation and provided written informed consent. They were assured of anonymity and informed about the purpose of the research, how their data would be used, and that no risks were associated with their involvement. Among the 104 participants, only one was a minor (17 years old), for whom guardian consent was obtained.

REFERENCES

[1] M. L. Maher, B. Skow, and A. Cicognani, "Designing the virtual campus," Design Studies, vol. 20, no. 4, pp. 319–342, Jul. 1999, doi: 10.1016/S0142-694X(98)00043-X.

[2] B. C. Nelson and D. J. Ketelhut, "Scientific inquiry in educational multi-user virtual environments," Educational Psychology Review, vol. 19, no. 3, pp. 265–283, Sep. 2007, doi: 10.1007/s10648-007-9048-1.

[3] N. Pellas, I. Kazanidis, N. Konstantinou, and G. Georgiou, "Exploring the educational potential of three-dimensional multi-user virtual worlds for STEM education: A mixed-method systematic literature review," Education and Information Technologies, vol. 22, no. 5, pp. 2235–2279, Sep. 2017, doi: 10.1007/s10639-016-9537-2.

[4] J. C. Chen, "The crossroads of English language learners, task-based instruction, and 3D multi-user virtual learning in Second Life," Computers & Education, vol. 102, pp. 152–171, Nov. 2016, doi: 10.1016/j.compedu.2016.08.004.

[5] I.-H. Cho, J.-H. Yeo, G.-H. Hwang, and H.-H. Yang, "Impact of a virtual environment on the learning effectiveness, motivation, cognitive load, and group self-efficacy of elementary school students in collaborative learning," Educational Technology Research and Development, vol. 70, no. 6, pp. 2145–2169, Dec. 2022, doi: 10.1007/s11423-022-10159-z.

[6] P. Li, J. Legault, A. Klippel, and J. Zhao, "Virtual reality for student learning: Understanding individual differences," Human Behaviour and Brain, vol. 1, no. 1, pp. 28–36, Apr. 2020, doi: 10.37716/hbab.2020010105.

[7] H. A. T. Van Limpt-Broers, M. Postma, E. Van Weelden, S. Pratesi, and M. M. Louwerse, "Neurophysiological evidence for the overview effect: A virtual reality journey into space," Virtual Reality, vol. 28, no. 3, p. 140, Jul. 2024, doi: 10.1007/s10055-024-01035-7.

[8] M. Akgün and B. Atici, "The effects of immersive virtual reality environments on students' academic achievement: A meta-analytical and meta-thematic study," Participatory Educational Research, vol. 9, no. 3, pp. 111–131, May 2022, doi: 10.17275/per.22.57.9.3.

[9] J. Conesa, F. J. Mula, K. A. Bartlett, F. Naya, and M. Contero, "The influence of immersive and collaborative virtual environments in improving spatial skills," Applied Sciences, vol. 13, no. 14, p. 8426, Jul. 2023, doi: 10.3390/app13148426.

[10] R. Kaliisa et al., "A topical review of research in computer-supported collaborative learning: Questions and possibilities," Computers & Education, vol. 228, p. 105246, Apr. 2025, doi: 10.1016/j.compedu.2025.105246.

[11] G. Makransky and G. B. Petersen, "The theory of immersive collaborative learning (TICOL)," Educational Psychology Review, vol. 35, no. 4, p. 103, Dec. 2023, doi: 10.1007/s10648-023-09822-5.

[12] J. Zambrano R., F. Kirschner, J. Sweller, and P. A. Kirschner, "Effects of prior knowledge on collaborative and individual learning," Learning and Instruction, vol. 63, p. 101214, Oct. 2019, doi: 10.1016/j.learninstruc.2019.05.011.

[13] A. L. V. Gaad, "The effects of online collaborative learning (OCL) on student achievement and engagement," IAFOR Journal of Education, vol. 10, no. 3, pp. 31–49, Dec. 2022, doi: 10.22492/ije.10.3.02.

[14] V. T. S. Law, H. H. L. Yee, T. K. C. Ng, and B. Y. F. Fong, "Transition from traditional to online learning in Hong Kong tertiary educational institutions during COVID-19 pandemic," Technology, Knowledge and Learning, vol. 28, no. 3, pp. 1425–1441, Sep. 2023, doi: 10.1007/s10758-022-09603-z.

[15] F. Wehrmann and R. Zender, "A systematic review of cooperation in multi-user virtual reality learning environments," Journal of Computer Assisted Learning, vol. 41, no. 5, p. e70112, Oct. 2025, doi: 10.1111/jcal.70112.

[16] W.-Y. Hwang and S.-S. Hu, "Analysis of peer learning behaviors using multiple representations in virtual reality and their impacts on geometry problem solving," Computers & Education, vol. 62, pp. 308–319, Mar. 2013, doi: 10.1016/j.compedu.2012.10.005.

[17] C.-M. Chen and M.-Y. Huang, "Enhancing programming learning performance through a Jigsaw collaborative learning method in a metaverse virtual space," International Journal of STEM Education, vol. 11, no. 1, p. 36, Aug. 2024, doi: 10.1186/s40594-024-00495-2.

[18] M. Sedlák, Č. Šašinka, Z. Stachoň, J. Chmelík, and M. Doležal, "Collaborative and individual learning of geography in immersive virtual reality: An effectiveness study," PLOS ONE, vol. 17, no. 10, p. e0276267, Oct. 2022, doi: 10.1371/journal.pone.0276267.

[19] M. M. Thompson, A. Wang, D. Roy, and E. Klopfer, "Authenticity, interactivity, and collaboration in VR learning games," Frontiers in Robotics and AI, vol. 5, p. 133, Dec. 2018, doi: 10.3389/frobt.2018.00133.

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

[20] M. Ashoori, Z. Shen, C. Miao, and L. Peyton, "Pedagogical agents for personalized multi-user virtual environments," International Journal of Engineering Education, vol. 25, no. 4, pp. 772–776, 2009.

[21] I. Hidrogo, D. Zambrano, M. Hernandez-de-Menendez, and R. Morales-Menendez, "Mostla for engineering education: Part 1 initial results," International Journal on Interactive Design and Manufacturing, vol. 14, no. 4, pp. 1429–1441, Dec. 2020, doi: 10.1007/s12008-020-00730-4.

[22] X. Huang and F. Macgilchrist, "A virtual classroom map-based immersive VR learning approach to fostering collaborative learning," Computers & Education: X Reality, vol. 5, p. 100088, Dec. 2024, doi: 10.1016/j.cexr.2024.100088.

[23] C. M. Stracke et al., "Immersive virtual reality in higher education: A systematic review of the scientific literature," Virtual Reality, vol. 29, no. 2, p. 64, Apr. 2025, doi: 10.1007/s10055-025-01136-x.

[24] M. Coban, Y. I. Bolat, and I. Goksu, "The potential of immersive virtual reality to enhance learning: A meta-analysis," Educational Research Review, vol. 36, p. 100452, Jun. 2022, doi: 10.1016/j.edurev.2022.100452.

[25] K. Nakano, N. Isoyama, D. Monteiro, N. Sakata, K. Kiyokawa, and T. Narumi, "Head-mounted display with increased downward field of view improves presence and sense of self-location," IEEE Transactions on Visualization and Computer Graphics, vol. 27, no. 11, pp. 4204–4214, Nov. 2021, doi: 10.1109/TVCG.2021.3106513.

[26] Y. Cao, G.-W. Ng, and S.-S. Ye, "Design and evaluation for immersive virtual reality learning environment: A systematic literature review," Sustainability, vol. 15, no. 3, p. 1964, Jan. 2023, doi: 10.3390/su15031964.

[27] T. T. De Back, A. M. Tinga, and M. M. Louwerse, "Learning in immersed collaborative virtual environments: Design and implementation," Interactive Learning Environments, vol. 31, no. 8, pp. 5364–5382, Nov. 2023, doi: 10.1080/10494820.2021.2006238.

[28] R. E. Mayer, "The past, present, and future of the cognitive theory of multimedia learning," Educational Psychology Review, vol. 36, no. 1, p. 8, Mar. 2024, doi: 10.1007/s10648-023-09842-1.

[29] L. Paulsen, S. Dau, and J. Davidsen, "Designing for collaborative learning in immersive virtual reality: A systematic literature review," Virtual Reality, vol. 28, no. 1, p. 63, Mar. 2024, doi: 10.1007/s10055-024-00975-4.

[30] L. Paulsen and J. Davidsen, "Activity-based collaborative virtual reality: Conceptualising immersive virtual reality for collaborative learning," International Journal of Computer-Supported Collaborative Learning, vol. 20, no. 3, pp. 317–341, Sep. 2025, doi: 10.1007/s11412-025-09446-7.

[31] J. Birt and N. Vasilevski, "Comparison of single and multiuser immersive mobile virtual reality usability in construction education," Educational Technology & Society, vol. 24, no. 2, pp. 93–106, 2021, doi: 10.2307/27004934.

[32] S.-J. Lu, Y.-C. Lin, K. H. Tan, and Y.-C. Liu, "Revolutionizing elementary disaster prevention education and training via augmented reality-enhanced collaborative learning," International Journal of Engineering Business Management, vol. 14, p. 18479790211067345, Nov. 2022, doi: 10.1177/18479790211067345.

[33] L. Xia, L. Wang, and C. Huang, "Implementing a social presence-based teaching strategy in online lecture learning," European Journal of Investigation in Health, Psychology and Education, vol. 14, no. 9, pp. 2580–2597, Sep. 2024, doi: 10.3390/ejihpe14090170.

[34] E. Anghel, "Let's do this together: Do the quantity and the quality of collaborative learning predict achievement among college students?," Active Learning in Higher Education, vol. 25, no. 2, pp. 227–241, Jul. 2024, doi: 10.1177/14697874221120783.

[35] J. Wang, "The comparison between architecture students' peer learning in informal situations within physical and virtual environments during the COVID-19 pandemic," Indoor and Built Environment, vol. 32, no. 10, pp. 2064–2082, Dec. 2023, doi: 10.1177/1420326X231193576.

[36] M. Dindar, L. Ren, and H. Järvenoja, "An experimental study on the effects of gamified cooperation and competition on English vocabulary learning," British Journal of Educational Technology, vol. 52, no. 1, pp. 142–159, Jan. 2021, doi: 10.1111/bjet.12977.

[37] J. Parong and R. E. Mayer, "Cognitive and affective processes for learning science in immersive virtual reality," Journal of Computer Assisted Learning, vol. 37, no. 1, pp. 226–241, Feb. 2021, doi: 10.1111/jcal.12482.

[38] Y. Tao, G. Zhang, D. Zhang, F. Wang, Y. Zhou, and T. Xu, "Exploring persona characteristics in learning: A review study of pedagogical agents," Procedia Computer Science, vol. 201, pp. 87–94, 2022, doi: 10.1016/j.procs.2022.03.014.

[39] M. Lachmair, M. H. Fischer, and P. Gerjets, "Action-control mappings of interfaces in virtual reality: A study of embodied interaction," Frontiers in Virtual Reality, vol. 3, p. 976849, Nov. 2022, doi: 10.3389/frvir.2022.976849.

[40] O. Chernikova et al., "Personalization through adaptivity or adaptability? A meta-analysis on simulation-based learning in higher education," Educational Research Review, vol. 46, p. 100662, Feb. 2025, doi: 10.1016/j.edurev.2024.100662.

[41] G. Makransky, L. Lilleholt, and A. Aaby, "Development and validation of the Multimodal Presence Scale for virtual reality environments: A confirmatory factor analysis and item response theory approach," Computers in Human Behavior, vol. 72, pp. 276–285, Jul. 2017, doi: 10.1016/j.chb.2017.02.066.

[42] D. Roth and M. E. Latoschik, "Construction of the Virtual Embodiment Questionnaire (VEQ)," IEEE Transactions on Visualization and Computer Graphics, vol. 26, no. 12, pp. 3546–3556, Dec. 2020, doi: 10.1109/TVCG.2020.3023603.

[43] J.-Y. Jian, A. M. Bisantz, and C. G. Drury, "Foundations for an empirically determined scale of

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

trust in automated systems," International Journal of Cognitive Ergonomics, vol. 4, no. 1, pp. 53–71, Mar. 2000, doi: 10.1207/s15327566ijce0401_04.

[44] G.-J. Hwang, L.-H. Yang, and S.-Y. Wang, "A concept map-embedded educational computer game for improving students' learning performance in natural science courses," Computers & Education, vol. 69, pp. 121–130, Nov. 2013, doi: 10.1016/j.compedu.2013.07.008.

[45] E. Ai-Lim Lee, K. W. Wong, and C. C. Fung, "How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach," Computers & Education, vol. 55, no. 4, pp. 1424–1442, Dec. 2010, doi: 10.1016/j.compedu.2010.06.006.

[46] S. Y. Tadimalla and M. L. Maher, "AI literacy for all: Adjustable interdisciplinary socio-technical curriculum," in 2024 IEEE Frontiers in Education Conference (FIE), Oct. 2024, pp. 1–9, doi: 10.1109/FIE61694.2024.10893159.

[47] S.-C. Kong, W. M.-Y. Cheung, and O. Tsang, "Evaluating an artificial intelligence literacy programme for empowering and developing concepts, literacy and ethical awareness in senior secondary students," Education and Information Technologies, vol. 28, no. 4, pp. 4703–4724, Apr. 2023, doi: 10.1007/s10639-022-11408-7.

[48] T. Xu et al., "From recorded to AI-generated instructional videos: A comparison of learning performance and experience," British Journal of Educational Technology, vol. 56, no. 4, pp. 1463–1487, Jul. 2025, doi: 10.1111/bjet.13530.

[49] J. F. Hair, J. J. Risher, M. Sarstedt, and C. M. Ringle, "When to use and how to report the results of PLS-SEM," European Business Review, vol. 31, no. 1, pp. 2–24, Jan. 2019, doi: 10.1108/EBR-11-2018-0203.

[50] H.-J. So and T. A. Brush, "Student perceptions of collaborative learning, social presence and satisfaction in a blended learning environment: Relationships and critical factors," Computers & Education, vol. 51, no. 1, pp. 318–336, Aug. 2008, doi: 10.1016/j.compedu.2007.05.009.

[51] J. Cohen, "A coefficient of agreement for nominal scales," Educational and Psychological Measurement, vol. 20, no. 1, pp. 37–46, 1960, doi: 10.1177/001316446002000104.

[52] R. J. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," Biometrics, vol. 33, no. 1, pp. 159–174, Mar. 1977, doi: 10.2307/2529310.

[53] D. A. Kenny, D. A. Kashy, and W. L. Cook, *Dyadic Data Analysis*. New York, NY, USA: Guilford Press, 2006.

[54] J. Cohen, "A power primer," Psychological Bulletin, vol. 112, no. 1, pp. 155–159, 1992, doi: 10.1037/0033-2909.112.1.155.

[55] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences,* 2nd ed. New York, NY: Psychology Press, 2009.

[56] J. Sigudla and J. E. Maritz, "Exploratory factor analysis of constructs used for investigating research uptake for public healthcare practice and policy in a resource-limited setting, South Africa," BMC Health Services Research, vol. 23, no. 1, p. 1423, Dec. 2023, doi: 10.1186/s12913-023-10165-8.

[57] T. K. Koo and M. Y. Li, "A guideline of selecting and reporting intraclass correlation coefficients for reliability research," Journal of Chiropractic Medicine, vol. 15, no. 2, pp. 155–163, Jun. 2016, doi: 10.1016/j.jcm.2016.02.012.

[58] A. Åkerfeldt, N. Bergdahl, and S. Hrastinski, "Adult learners' perceptions of distance education," Journal of Adult and Continuing Education, vol. 30, no. 1, pp. 248–266, May 2024, doi: 10.1177/14779714231200422.

[59] R. Ratan et al., "Time matters in VR: Students benefit from longer VR class duration, but certain outcomes decline after 45 minutes, with large individual variance," Computers & Education, vol. 235, p. 105328, Oct. 2025, doi: 10.1016/j.compedu.2025.105328.

[60] R. Sterna and K. Zibrek, "Psychology in virtual reality: Toward a validated measure of social presence," Frontiers in Psychology, vol. 12, p. 705448, Oct. 2021, doi: 10.3389/fpsyg.2021.705448.

[61] N. Aburumman, M. Gillies, J. A. Ward, and A. F. de C. Hamilton, "Nonverbal communication in virtual reality: Nodding as a social signal in virtual interactions," International Journal of Human-Computer Studies, vol. 164, p. 102819, Aug. 2022, doi: 10.1016/j.ijhcs.2022.102819.

[62] L. Foglia and R. A. Wilson, "Embodied cognition," WIREs Cognitive Science, vol. 4, no. 3, pp. 319–325, May 2013, doi: 10.1002/wcs.1226.

[63] D. Maloney, G. Freeman, and D. Y. Wohn, "'Talking without a voice': Understanding non-verbal communication in social virtual reality," Proceedings of the ACM on Human-Computer Interaction, vol. 4, no. CSCW2, pp. 1–25, Oct. 2020, doi: 10.1145/3415246.

[64] F. B. Schneider, Ed., *Trust in Cyberspace*. Washington, DC, USA: National Academy Press, 1999.

[65] P. Vlachogianni and N. Tselios, "Perceived usability evaluation of educational technology using the Post-Study System Usability Questionnaire (PSSUQ): A systematic review," Sustainability, vol. 15, no. 17, p. 12954, Aug. 2023, doi: 10.3390/su151712954.

[66] G. Makransky, T. S. Terkildsen, and R. E. Mayer, "Adding immersive virtual reality to a science lab simulation causes more presence but less learning," Learning and Instruction, vol. 60, pp. 225–236, Apr. 2019, doi: 10.1016/j.learninstruc.2017.12.007.

[67] I. Altman and D. A. Taylor, *Social Penetration: The Development of Interpersonal Relationship*s. New York, NY, USA: Holt, Rinehart and Winston, Inc., 1973.

[68] A. C. Edmondson and D. P. Bransby, "Psychological safety comes of age: Observed themes in an established literature," Annual Review of Organizational Psychology and Organizational Behavior, vol. 10, pp. 55–78, Jan. 2023, doi: 10.1146/annurev-orgpsych-120920-055217.

[69] J. Lim and M. Lee, "The buffering effects of using avatars in synchronous video conference-based online learning on students' concerns about interaction and

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2026.3744604

negative emotions," Education and Information Technologies, vol. 29, no. 12, pp. 16073–16096, 2024, doi: 10.1007/s10639-024-12508-2.

[70] M. R. Leary, *Self-Presentation: Impression Management and Interpersonal Behavior*, 1st ed. Routledge, 2019. doi: 10.4324/9780429497384.

[71] A. V. Oje, N. J. Hunsu, and L. Fiorella, "A systematic review of evidence-based design and pedagogical principles in educational virtual reality environments," Educational Research Review, vol. 47, p. 100676, May 2025, doi: 10.1016/j.edurev.2025.100676.

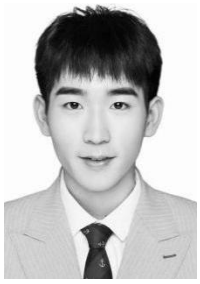
[72] C. Lazou, A. Tsinakos, and I. Kazanidis, "A rubric for peer evaluation of multi-user virtual environments for education and training," Information, vol. 16, no. 3, p. 174, Feb. 2025, doi: 10.3390/info16030174.

Yueyao Qu received his master's degree from Shaanxi Normal University, Xi'an, China, in 2026. Currently working at Chengdu No.7 Middle School Junior High School.

His research interests include educational technology and virtual reality in education.

Yun Zhou (Member, IEEE) received the Ph.D. degree in computer science from Ecole Centrale de Lyon, Ecully, France, in 2012.

She is currently a full professor of education at Shaanxi Normal University, Xi'an, China. Her research interests include educational virtual reality, brain-computer interfaces, and innovative interaction technologies for education.



Jin Shu received the B.S. degree in computer science from the School of Computer Science, Shaanxi Normal University, Xi'an, China.

He is currently working toward the master's degree with the Faculty of Education, Shaanxi Normal University, Xi'an, in 2024. His research interests include artificial intelligence in education, educational virtual reality, and spatial ability.

Junjun Zhai received her master's degree from Shaanxi Normal University, Xi'an, China.

Her research interests include virtual reality technology and educational technology.

Jie Bai received the Ph.D. degree in early childhood education from Shaanxi Normal University, Xi'an, China, in 2024.

She is currently an assistant researcher at Shaanxi Normal University. Her research interests include generative artificial intelligence, early childhood reading development, and educational technology for emergent literacy.

Jialin Yu received an M.S. degree in Modern Educational Technology from Nanchang University, Nanchang, China, in 2014.

She is currently an Experimental Engineer at the Faculty of Education, Shaanxi Normal University, Xi'an, China. Her research interests include virtual reality in education and digital learning resource development.

Tao Xu (Member, IEEE) received the Ph.D. degree in computer science from Ecole Centrale de Lyon, Ecully, France, in 2013.

He is currently a full professor at Northwestern Polytechnical University, Xi'an, China. His research interests include educational virtual reality and artificial intelligence in education.

Authorized licensed use limited to: Shaanxi Normal University. Downloaded on October 26,2026 at 15:48:26 UTC from IEEE Xplore. Restrictions apply.

© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.