Virtual Reality Modes in Hotel English Instruction: A Mixed-methods Study of Chinese University Students' Spatial Presence

Huan Meng and Mageswaran Sanmugam *

Centre for Instructional Technology and Multimedia, Universiti Sains Malaysia, Penang, Malaysia
Email: shanga202202@163.com (H.M.); mageswaran@usm.my (M.S.)
*Corresponding author

Manuscript received December 17, 2025; revised January 6, 2026; accepted February 24, 2026; published August 12, 2026

***Abstract*—Although Virtual Reality (VR) is widely regarded as an effective tool for enhancing experiential learning, its role in fostering spatial presence in English for Specific Purposes (ESP), particularly in a hotel English course, remains underexplored. This study examines whether Immersive VR (IVR) elicits stronger spatial presence than Desktop VR (DVR) and how these differences are reflected in learners' experiences of self-location and possible actions. To systematically investigate these relationships, a dual-mode VR hotel English training system was developed, with identical content, tasks, and instructional design across conditions, differing only in interaction modality. An explanatory sequential mixed-methods design was employed (QUANTITATIVE → qualitative = explanation). In the quantitative phase, 144 undergraduate students were assigned to IVR or DVR conditions in a quasi-experimental study. In the qualitative phase, 8 participants were interviewed to further explore their experiences. Quantitative results showed that students in the IVR condition reported significantly higher spatial presence than those in the DVR condition. Qualitative findings indicated that IVR enhanced learners' sense of self-location and possible actions through greater immersion and interactivity, although some participants reported discomfort and occasional difficulty concentrating. DVR, while less immersive, was perceived as more comfortable and easier to use. By integrating spatial presence theory with ESP pedagogy, this study provides empirical evidence for selecting appropriate VR modalities for hotel English instruction and offers a more explanatory framework to narrow the gap between students' conversational competence and the communicative demands of the hospitality industry.**

***Keywords*—virtual reality modes, Immersive Virtual Reality (IVR), Desktop Virtual Reality (DVR), hotel English, spatial presence**

I. Introduction

In pursuit of greater international competitiveness, China has implemented policies to expand global engagement, which has driven rapid economic growth and boosted the tourism industry [1]. As a key service-oriented sector, the hotel industry now faces increasingly high standards, requiring frontline staff to possess both professional service skills and English communication abilities [2].

To meet these demands, Hotel English, a branch of English for Specific Purposes (ESP), has been incorporated into hotel management programs in Chinese universities since the 1990s, with an emphasis on industry-specific terminology and situational dialogues [3]. However, despite ongoing curriculum reform, Hotel English instruction continues to face persistent challenges. Only 28.57% of employers consider hotel management graduates proficient in English communication, whereas 71.43% believe their language skills fall short of workplace requirements [4]. This discrepancy indicates that existing instructional approaches are limited in preparing students for the interaction-intensive tasks encountered in authentic hotel service contexts. It highlights a significant mismatch between graduates' English proficiency and the actual demands of the hospitality industry, underscoring the urgent need for innovative instructional approaches.

Educational technology offers potential solutions to these challenges in language instruction. Computer-Assisted Language Learning (CALL) has evolved to include multimedia and interactive platforms, but its capacity to create immersive, contextualized learning remains limited [5]. Virtual Reality (VR), as an advanced extension of CALL, offers learners immersive and interactive environments that simulate authentic language use scenarios, thereby enabling them to practice dialogue skills in real-life contexts [6]. In VR-based language learning, spatial presence refers to learners' subjective experience of being situated within a virtual environment and constitutes a core experiential mechanism underpinning the pedagogical effectiveness of VR [7]. By enhancing spatial presence, learners can perceive themselves as active participants in service interaction scenarios, enabling them to practice context-appropriate dialogue and professional behaviors. This immersive experience not only increases learners' engagement and task involvement but also helps bridge the gap between classroom-based Hotel English training and the dialogue demands of the hospitality industry. Consequently, students are better prepared to meet real-world workplace dialogue requirements and professional expectations. VR-based learning environments are typically classified as Immersive Virtual Reality (IVR) and Desktop Virtual Reality (DVR). IVR, typically implemented through Head-Mounted Displays (HMDs), Cave Automatic Virtual Environment (CAVE) systems, or projection-based platforms, offers high levels of immersion and interactivity, whereas DVR relies on conventional screens such as computers or mobile devices [8].

Recent research has highlighted the potential of VR to enhance language learning outcomes, including vocabulary acquisition, learner engagement, and sense of presence [9–12]. However, empirical evidence comparing how different VR modes influence learners' spatial presence remains limited, particularly in the context of English for Specific Purposes (ESP), such as hotel service dialogues.

Unlike previous studies, which typically examine a single VR mode or focus on general language learning contexts, this study adopts a comparative approach by examining Immersive Virtual Reality (IVR) and Desktop Virtual Reality (DVR), with spatial presence as a focal construct, within a Hotel English (ESP) instructional environment.

This study aims to investigate the differential effects of IVR and DVR on the spatial presence of Chinese university students' during Hotel English dialogue learning. It seeks to determine whether IVR elicits higher spatial presence than DVR and how these differences manifest in learners' experiences of self-location and possible actions within the virtual environment. Using a mixed-methods design that combines a quasi-experimental pre-test and post-test approach with semi-structured interviews, the study provides empirical insights into how VR mode selection influences ESP instruction and clarifies how spatial presence functions as a mechanism enabling VR to bridge the gap between classroom-based Hotel English training and the dialogue demands of the hospitality industry.

II. Literature Review

A. Hotel Management and Language Education

The hotel industry is a highly globalized and service-oriented sector in which effective English dialogue is essential for ensuring service quality and customer satisfaction [13, 14]. Within hospitality education, Hotel English has become an important branch of English for Specific Purposes (ESP), focusing on industry-specific vocabulary, situational dialogues, and the dialogue competence required in real workplace settings. Unlike General English, Hotel English emphasizes task-based and context-specific language use, such as front desk reception, complaint handling, telephone dialogues, and cross-cultural interaction, reflecting the functional and pragmatic demands of hotel service encounters [15].

Despite its practical orientation, traditional hotel English instruction continues to face challenges in developing learners' authentic communicative ability [16]. Common limitations include insufficient exposure to realistic service contexts, limited opportunities for authentic spoken interaction, and learners' anxiety when performing role-based dialogues [17]. Although recent studies have begun to explore the use of digital technologies, such as virtual reality, to enhance instructional authenticity, learner-level experiential factors have received limited investigation [18]. The way in which instructional design can foster learners' spatial presence and support effective hotel English dialogue learning has received limited systematic attention. This highlighting the need for pedagogical approaches that better align language instruction with the experiential demands of professional hospitality contexts.

B. Current Technological Solutions and Challenges

With the growing integration of digital technologies in hotel English instruction, various instructional approaches such as online learning, Information and Communications Technology (ICT), blended learning, Augmented Reality (AR), and flipped classrooms have been adopted to enhance instructional flexibility and learning outcomes. Previous studies have shown that online learning provides flexible access to resources, while learners often experience reduced engagement and limited interaction [19]. ICT can enhance opportunities for listening and speaking practice [20]; however, inadequate infrastructure and insufficient teacher training may constrain its effectiveness [21]. Blended learning combines online and offline instruction to support skill development; however, technical difficulties and poorly structured courses can reduce learning quality [22]. AR can create more immersive and interactive learning scenarios [23]; however, device compatibility, technical complexity, and high costs may hinder students' sense of presence [24]. Flipped classrooms promote self-directed learning and active participation [25]; however, high demands on time management and insufficient guidance can reduce students' engagement [26].

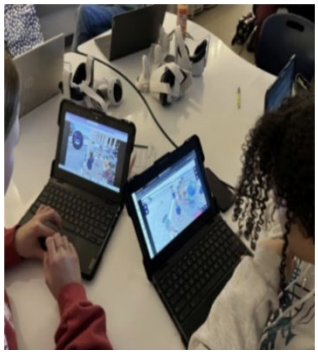
Overall, these technological approaches provide innovative opportunities for enhance teaching flexibility and learning outcomes in hotel English instruction, but they still face some limitations and challenges. In addition, their specific influence on students' spatial presence, a key factor in virtual learning experiences, has yet to be systematically compared. Therefore, this study investigates the effects of different VR modes on students' spatial presence and aims to provide insights into the practical application of VR in hotel English instruction.

C. Comparing Immersive and Desktop VR Applications in Language Learning

This study examines the effectiveness of IVR and DVR in language learning, particularly in the context of hotel English instruction. To systematically present the differences and advantages of the two approaches across key dimensions, Table 1 provides a comparative analysis of IVR and DVR. The aim is to provide foundational support for subsequent theoretical discussions and empirical investigations.

Table 1. Comparison of immersive and Desktop VR in language learning

Comparison Dimension	IVR	DVR	Reference
Visual Presentation	Uses Head-Mounted Display (HMD) supporting stereoscopic vision	Uses desktop monitor or mobile device	[27]
Sense of Immersion	High immersion with strong spatial presence	Lower immersion, experience is more observational	[28]
Interactivity	High interactivity through head and hand movements	Lower interactivity; relies on mouse and keyboard operations	[29]
Virtual Dimension	3D environment	2D interface	[27]
Sense of Presence	Stronger presence, immersive experience simulates authentic communication	Weaker sense of presence, more suitable for short-term experiences	[28]
Technical Cost	Higher cost, expensive equipment, and maintenance; limited accessibility	Lower cost, minimal barriers to use, more easily scalable	[30]

Physical Effects	Extended use may cause motion sickness, dizziness	Almost no physical discomfort	[31]
Space and Time Limitations	Requires designated space and equipment; limited by physical environment	Accessible anytime and anywhere via computers or mobile devices	[32]
Scope of Application	Effective for in-depth, small-scale instruction; limited by space and equipment	Widely applicable in education, easy for large-scale dissemination	[33]
Technical Adaptability	Longer adaptation period, more complex equipment, higher technical threshold; less beginner-friendly	Easy to operate, no equipment needed, quick to adapt, suitable for VR beginners	[34]
Instructional Sustainability	Resource-intensive; depends on institutional technology and funding	Suitable for large-scale implementation; high sustainability	[33]
Examples			—

The comparative analysis shows that IVR provides greater immersion and interactivity, which offers a stronger sense of spatial presence and more authentic dialogue experiences. However, it requires more complex equipment, dedicated space, and higher costs. In contrast, DVR is more accessible, easier to operate, and more cost-effective, making it suitable for large-scale instruction, although it provides lower immersion and a weaker sense of presence. These differences suggest that the two VR modes may stimulate learners' spatial presence to varying degrees depending on the instructional context. Therefore, investigating their application in hotel English instruction can provide insights into how VR design influences students' spatial presence and learning experience.

D. Previous Research with Spatial presence

Spatial presence refers to the perception of "being there" within a mediated environment, where users feel physically present and able to interact with the objects and people represented in the virtual or media-generated space [35]. High levels of immersion, such as those provided by stereoscopic vision and body tracking, can enhance spatial presence, although less immersive media can also elicit similar "being there" experiences [36]. Previous research has distinguished between physical and psychological immersion, emphasizing that spatial presence depends on both device-mediated sensory engagement and subjective involvement [37]. In VR learning, Head-Mounted Devices (HMDs) typically provide higher immersion than desktop monitors, which can increase spatial presence but may also raise extraneous cognitive load [38].

In the context of English for Specific Purposes (ESP), such as Hotel English courses, spatial presence is particularly important because it affects learners' sense of "being there" in simulated professional scenarios [39]. Professional scenario simulations require learners to engage in realistic interactions, including guest communication, service procedures, and cross-cultural exchanges [40]. A higher sense of spatial presence can make these experiences more authentic, supporting language output and task-based learning [41].

Several self-report instruments have been developed to assess spatial presence, including the Presence Questionnaire (PQ) [42], the Independent Television Commission-Sense of Presence Inventory (ITC-SOPI) [43], the Igroup Presence Questionnaire (IPQ) [42], and the Temple Presence Inventory (TPI) [44].

Despite their widespread use, these instruments have certain limitations. Only a few scales, such as the IPQ and ITC-SOPI, have undergone extensive psychometric testing, which complicates the comprehensive evaluation of other measures. Moreover, some subscales may capture methodological artifacts or peripheral aspects (e.g., "negative effects" in ITC-SOPI, "interface quality" in PQ) rather than core components of spatial presence [42].

To address these limitations, the Spatial Presence Experience Scale (SPES) was developed [45]. Derived from the process model of spatial presence [46], the SPES conceptualizes spatial presence as a two-dimensional construct, consisting self-location and possible actions. It provides a concise and practical tool for VR research. Previous studies have demonstrated that the SPES possesses good psychometric properties, supporting its suitability for assessing learners' subjective spatial presence.

E. Theories and Models of Research

This study based primarily on Spatial Presence Theory (SPT) [46] and Alessi and Trollip's instructional design model [47]. SPT, derived from the Two-Level Model, provides the theoretical foundation for conceptualizing spatial presence and explaining how learners experience immersion in virtual environments. It also justifies the use of the Spatial Presence Experience Scale (SPES) as an established instrument for measuring this construct. In addition, Alessi and Trollip's instructional design model offers a structured and systematic framework, guiding the design and development of the dual-mode VR Hotel English training system used in this study.

1) Spatial Presence Theory (SPT)

Spatial Presence Theory (SPT) describes the psychological experience in which individuals feel physically present in a mediated environment, despite being aware of its artificial nature. Witmer [48] defined presence as the subjective sensation of "being there," laying the foundation for

subsequent theoretical elaborations. Subsequent scholarship generally agrees that spatial presence is a subjective cognitive experience arising from perceptual, attentional, and cognitive processes involved in spatial information processing [49]. In VR systems, these processes may occur automatically, leading individuals to experience presence even when they fully recognize the mediated nature of the environment.

The theoretical development of spatial presence includes several influential models that emphasize different underlying mechanisms. Draper [50] proposed an attention-based model, arguing that spatial presence emerges when attentional is fully allocated to the mediated environment. Steuer [51] highlighted sensory richness, positing that vivid perceptual stimulation is central to telepresence. Slater [52] conceptualized presence as dependent on "doing there," suggesting that active interaction and sensorimotor contingencies are primary determination. Schubert [53] extended this approach through an embodied cognition framework, emphasizing the meaningful and intuitive bodily engagement. Although these perspectives differ, they converge on several points: spatial presence requires focused attention, sensorimotor coupling enhances the illusion of "being there," and bodily interaction with consistent feedback is critical. These insights informed the development of a more comprehensive and integrative model, namely Two-Level Model [46], which is now widely adopted in presence research.

The Two-Level Model of spatial presence [46] consolidates earlier fragmented explanations by conceptualizing spatial presence as the outcome of two core psychological components: self-location, referring to the perceptual experience of being situated within the mediated spatial environment, and possible actions, referring to the perception of meaningful opportunities for action. When both components are activated, individuals experience a binary state of spatial presence, feeling located "inside" the mediated space and capable of acting within it. This model integrates prior work on attention [50], sensory processing [51], cognitive Gestalt formation [52], and embodied interaction [49]. Its strength lies in offering a systematic and measurable explanation of how spatial presence emerges, making it especially suitable for experimental and applied contexts. Importantly, it also serves as the theoretical foundation for the Spatial Presence Experience Scale (SPES), which operationalizes presence through items reflecting self-location and possible actions. The SPES has been empirically validated across various mediated environments [54], confirming strong reliability and construct validity. Given its integrative structure, empirical robustness, and direct correspondence with SPES, Two-Level Model [46] provides an appropriate theoretical basis for this study. It is specifically adopted to conceptualize and measure learners' spatial presence in IVR and DVR learning environments, ensuring that the psychological experience of "being there" and perceiving meaningful action opportunities is systematically captured and empirically validated. This framework, in turn, allows the study to examine how the design of the 'dual-mode VR hotel English training system' impacts learners' spatial presence hotel English course study.

2) Alessi and Trollip's Instructional Design (ATID) model

The Dual-Mode VR Hotel English Training System in this study was developed based on Alessi and Trollip's instructional design model [47]. The selection of this model is primarily based on three reasons: First, it is specifically designed for beginners and instructional developers, characterized by a clear structure and widespread applicability across various disciplines, including language learning. Second, although the model has been in use for over 20 years, its sub-stages offer a high degree of flexibility, allowing adjustments in response to advancements in interactive multimedia technology to meet the needs of IVR and DVR language learning environments. Lastly, the model has been widely applied in educational technology and system design research, demonstrating its guiding value [55]. Based on these factors, this study adopts the model to develop the dual-mode VR hotel English training system, aiming to optimize the learning experience and enhance learners' language proficiency. The model integrates the essential steps involved in the planning, design, and development stages of instructional system implementation. To clearly illustrate how this model is applied in the present study, Table 2 presents a mapping of the core tasks in each phase and their corresponding implementation in the research.

Table 2. The three phases of the Alessi & Trollip instructional design model and their application in this study

Phases	Core Tasks	Application in This Study
Planning	- Define the scope - Identify learner characteristics - Establish the constraints - Determine and collect resources	• Define the instructional goals and scope of the hotel English course • Identify the participants as students in a Northern Chinese city and analyze their English proficiency and VR experience • Consider practical constraints such as equipment availability, scheduling, and content difficulty • Collect VR instructional resources, English language corpora, technical support materials, etc.
Design	- Develop initial content ideas - Create storyboards - Prepare scripts - Prepare prototype	• Design task flow and contextual content aligned with hotel English scenarios • Use storyboards to plan task pathways and interaction logic • Write pragmatic scripts and English dialogue boxes • Develop an initial prototype of the VR hotel English instructional system
Development	- Prepare the text - Create the graphics - Produce audio and video - Assemble the pieces - Alpha testing - Make revisions - Beta testing	• Write instructional texts and convert them into scene scripts • Create visuals and character models for the virtual hotel • Record English voiceovers, dubbing, and environmental sound effects • Integrate all content into the VR system • Conduct a small-scale Alpha test and collect feedback • Make revisions and Adjust content based on the feedback • Conduct Beta testing before the formal experiment to validate the instructional process and technical stability

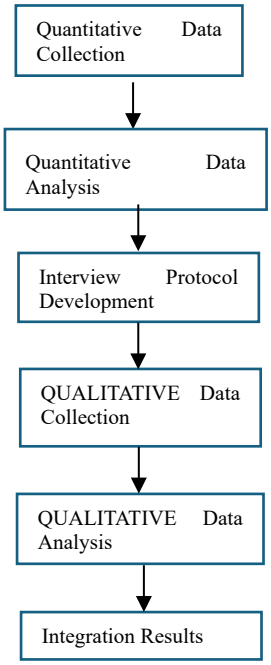
F. Research Gap

This study systematically reviewed 36 studies published between 2020 and 2025 on the application of IVR and DVR in English learning, aiming to summarize current trends, highlight key focus areas and limitations, and identify potential directions for future research. Existing studies have primarily focused on general English learning contexts and have rarely examined the intersection of English for Specific Purposes (ESP) and spatial presence. Research on professional scenario simulations for hotel English courses is limited, and the differential effects of IVR and DVR on students' spatial presence remain largely unexplored. To address this gap, the present study compares IVR and DVR in hotel English courses to examine their effects on spatial presence, thereby providing context-specific insights for VR-based ESP instruction.

III. Methodology

A. Research Design

This study employed the explanatory sequential mixed methods research design (QUANTITATIVE → qualitative = explanation), which included an initial quantitative quasi-experimental study and the followed by qualitative interview, with priority given to the quantitative phase. The qualitative findings were used to interpret the quantitative results and provide a deeper understanding of both significant and nonsignificant outcomes. A flowchart of the mixed methods sequential explanatory design procedures is shown in Fig. 1 (adapted from Ref. [56]).



Procedure	Product
• Quasi-experimental design • Spatial Presence Experience Scale (SPES)	• Numeric data
• Descriptive statistics: Compute mean, standard deviation, and frequency distribution for spatial presence. • ANCOVA: Compare post-test scores of IVR and DVR groups while controlling for pre-test scores. • SPSS Quan. software v.31.	• Tables of descriptive statistics (mean, SD, frequencies), • ANCOVA results (F-values, p-values, effect sizes). • interaction/moderation effects
• Purposefully selecting 2 participants from each group (N=8) based on VR modes and spatial presence (high/low) • Developing interview questions	• Cases (N=8) • Interview protocol
• Individual semi-structured interviews (30–45 min per participant) • Conduct interviews via in-person • Record and transcribe interviews • Collect observation notes and teaching material screenshots/videos	Text data (interview transcripts, documents, artifact description) • Image data (photographs)
• Coding and thematic analysis • Theme development • NVivo qualitative software	• Codes, themes, sub-themes • Visual representation of themes • tables summarizing patterns across participants
• Interpretation and explanation of the quantitative and qualitative results	• Discussion • Implications • Future research

Fig. 1. Explanatory sequential mixed-methods design procedures.

B. Quantitative Study

The quantitative research adopted a quasi-experimental design to evaluate the effects of IVR and DVR on students' spatial presence in hotel English instruction. A pre-test–post-test design was employed, with comparable intact classes purposively selected and assigned to two conditions: the IVR class served as the experimental group, and the DVR class served as the control group. Before the intervention, all students completed a pre-test measuring their initial level of spatial presence in hotel English learning. After a period of instruction using either IVR or DVR, students completed a post-test to examine the effectiveness of the two VR modes in enhancing spatial presence.

1) Population

For data collection, one university in northern China was selected through purposive sampling to ensure contextual relevance to hotel English instruction. The target population consisted of approximately 575 undergraduate students enrolled in the Hotel Management program at this university. From this population, participants were selected using simple random sampling, ensuring equal selection probability for all eligible students and thereby minimizing sampling bias.

2) Sample size and sampling technique

The sample size for this study was estimated using G*Power 3.1 software. Following the guidelines outline by Cohen [57], a medium effect size ($f = 0.25$), $\alpha = 0.05$, and a statistical power of 0.8 were selected. Based on these parameters, the required total sample size was approximately 128 participants. To account for potential missing or invalid data, an additional 10% oversampling was applied [58]. As a result, the final sample size was set at 36 participants per group, yielding a total of 144 participants. The detailed estimation parameters are presented in Table 3, and the G*Power output is shown in Fig. 2.

For the sampling procedure, purposive sampling was first used to select a university in northern China that offered the Hotel Management program and was suitable for implementing VR-based hotel English instruction. Within this university, participants were drawn from eligible

undergraduate students using simple random sampling, ensuring equal selection probability and minimizing selection bias. Finally, intact classes were used to assign students to the IVR and DVR groups, aligned with the existing class structure and teaching schedule while maintaining the integrity of the quasi-experimental pre-test–post-test design.

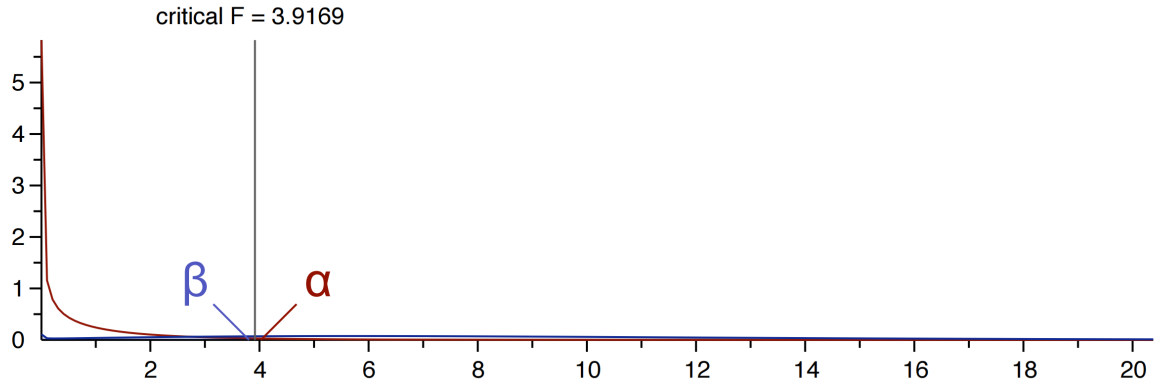


Fig. 2. Distribution plot of G*Power analysis.

Table 3. G*Power analysis for sample size estimation (2×2 factorial design)

Analysis	A priori: Compute required sample size –given a, power, and effect size	Value
Input:	Effect size f	= 0.25
	α err prob	= 0.05
	Power (1-β err prob)	= 0.8
	Numerator df	= 1
	Number of groups	= 2
Output:	Non-centrality parameter λ	= 8.0000000
	Critical F	= 3.9169322
	Denominator df	= 125
	Total sample size	= 128
	Actual power	= 0.8014112

Note: F tests—ANCOVA: Fixed effects, main effects and interactions.

3) Instrument development and validation

In the present study, the SPES was used to measure participants' spatial presence during the VR-based hotel English learning training system. The SPES is an 8-item self-report instrument with two subscales: self-location and possible actions. Participants responded on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree).

Before completing the SPES, participants engaged with the VR environment for at least 10 mins to ensure sufficient exposure. Minor wording adjustments were applied to align the SPES items with the IVR and DVR hotel English learning context. The scale was administered pre-test and post-test to capture changes in spatial presence.

The scale has demonstrated good psychometric properties, with Cronbach's alpha coefficients of $\alpha = 0.91$ for the self-location subscale and $\alpha = 0.84$ for the possible action subscale. The overall scale demonstrated an internal consistency of $\alpha = 0.797$. Confirmatory Factor Analysis (CFA) also supported the two-factor structure, indicating good construct validity.

4) Data collection

For the first, quantitative phase, a quasi-experimental design with pre-test and post-test procedures was employed to examine the effects of IVR and DVR instructional modes on students' hotel English learning. Data were collected using a structured procedure to ensure consistency and accuracy. The questionnaire was distributed via the WJX platform (www.wjx.cn) and completed on-site under the supervision to ensure data quality. Rest breaks and preparation time were provided to minimize fatigue and ensure high-quality responses. Before the final data collection, a pilot test with a small sample was conducted to assess item clarity and consistency, allowing minor adjustments before the formal administration.

5) Data analysis techniques

The quantitative data obtained in this study were analyzed using IBM Statistical Packages for Social Sciences (SPSS) version 31. Descriptive and inferential statistics were employed to summarize the data and test the research hypotheses. Before conducting ANCOVA, its assumptions were systematically evaluated to ensure analytical validity. Following the guidelines of Refs. [59, 60], three key assumptions were examined: (1) normality of residuals, (2) reliability of the covariate, and (3) homogeneity of variances. The detailed procedures for testing these assumptions, along with the corresponding results, are presented in Section IV (Results). If any of these assumptions had been violated, corrective measures such as variable transformation, exclusion of problematic covariates, or adoption of non-parametric parametric alternatives (e.g., Quade's test) would have been considered.

C. Qualitative Study

The qualitative component of this study employed semi-structured interviews and thematic analysis to explore students' spatial presence in hotel English courses across IVR and DVR environments. The qualitative design was grounded in Ref. [58], and the interview questions were developed based on the research question to capture a comprehensive understanding of students' learning experiences. The interview data were analyzed through thematic coding and theme extraction to identify how different VR modes shaped students' spatial presence. Finally, the qualitative findings were integrated with the quantitative results to provide a more holistic interpretation of the effects of VR modes in hotel English learning.

1) Participants

In the qualitative phase, 8 students were randomly selected from the quantitative sample, with 2 participants chosen from each group, to participate in semi-structured interviews. This

sampling strategy aimed to capture diverse perspectives on students' spatial presence in different VR modes of hotel English learning. The sample size was guided by the principle of data saturation. In relatively homogeneous samples, 6 to 12 participants are generally sufficient to achieve data saturation [61]. Given that the sample consisted of hotel management students from the same university with similar backgrounds, the selection of 8 participants was considered adequate to achieve meaningful saturation.

2) Semi-structured interview protocol

The qualitative data were collected using a semi-structured interview protocol developed by the researcher. The protocol consisted of two sections. The first section addressed ethical and procedural matters, which were covered in the informed consent form provided to participants. The second section consisted of the actual interview questions, designed to elicit participants' perspectives and experiences related to VR-based Hotel English learning and spatial presence.

The interview questions were aligned with the study's research questions and theoretical framework and were further informed by the established guidelines for qualitative interview design (e.g., [62]). Specifically, to ensure conceptual consistency with the spatial presence framework, the interviews focused on two core dimensions of students' spatial presence experiences: self-location and possible actions.

3) Data collection

In this study, the semi-structured interviews were conducted within one day after the completion of the experiment to ensure that participants could accurately recall their learning experiences. Each interview lasted approximately 30-45 min and was fully recorded under strict confidentiality.

Data were collected from multiple sources to ensure richness and depth for each participant's case. In-depth semi-structured interviews served as the primary data source, focusing on the spatial presence construction. Additional follow-up inquiries were conducted via email or online messaging to further explore emerging themes. To validate self-reported learning achievement, students' learning performance records, including pre-test and post-test scores, were examined. Elicitation materials, such as screenshots and recordings from VR learning sessions, were used to prompt reflection and capture participants' interactions and experiences. In addition, researcher observation notes taken during the VR sessions were also incorporated to provide contextual insights and further enrich the dataset. Data collection was conducted in two stages: a two-week pilot phase in mid-June 2025, followed by the main study in mid-October 2025.

All collected data were securely stored on a password-protected and encrypted personal computer accessible only to the researcher. Hard copies of the interview transcripts were kept in a locked cabinet to ensure confidentiality. All research procedures complied with the university's ethical guidelines, and the data will be permanently destroyed after the required retention period.

4) Data analysis

The qualitative data analysis in this study primarily employs thematic analysis to extract and summarize core themes from the interview data. First, the fundamental concepts and application principles of thematic analysis are outlined. Then, the specific steps and models of thematic analysis are explained in detail to ensure systematic and consistent data processing. Additionally, the supporting software tools used for analysis and methods for evaluating theme reliability and validity are introduced to enhance the credibility of the research findings.

IV. Results

A. Quantitative Findings

This section reports the quantitative findings of the study, including descriptive statistics for students' spatial presence scores and the results of ANCOVA examining the effects of VR mode on post-test spatial presence while controlling for pre-test scores.

1) Descriptive statistics

Table 4 presents the descriptive statistics for the pre-test and post-test results across the two VR modes (IVR and DVR). The tables summarize students' responses to eight items measuring their sense of spatial presence in the hotel English learning environment. For each item, the mean and standard deviation are reported for both the pre-test and post-test phases.

Table 4. Descriptive statistics for spatial presence in IVR and DVR conditions

VR Modes	Questions	Pre-test Mean	Pre-test SD	Post-test Mean	Post-test SD
IVR	1. I felt like I was actually there in the environment of the IVR hotel English learning.	4.78	1.124	5.14	0.990
	2. It seemed as though I actually took part in the activities of the IVR hotel English learning environment.	4.69	0.786	4.78	0.929
	3. It was as though my true location had shifted into the environment in the IVR hotel English learning.	4.92	0.937	5.28	0.914
	4. I felt as though I was physically present in the environment of the IVR hotel English learning.	4.67	0.956	5.03	0.941
	5. The objects in the IVR hotel English learning environment gave me the feeling that I could do things with them.	4.64	0.798	4.78	0.959
	6. I had the impression that I could be active in the environment of the IVR hotel English learning.	4.47	0.845	5.00	0.956
	7. I felt like I could move around among the objects in the IVR hotel English learning environment.	4.61	0.838	5.00	0.926
	8. It seemed to me that I could do whatever I wanted in the environment of the IVR hotel English learning.	4.28	0.741	4.56	0.652
DVR	1. I felt like I was actually there in the environment of the DVR hotel English learning.	3.97	1.207	5.17	1.254

2. It seemed as though I actually took part in the activities of the DVR hotel English learning environment.	4.78	0.760	5.50	1.000
3. It was as though my true location had shifted into the environment in the DVR hotel English learning.	4.64	0.639	5.31	0.920
4. I felt as though I was physically present in the environment of the DVR hotel English learning.	4.56	0.773	5.47	1.108
5. The objects in the DVR hotel English learning environment gave me the feeling that I could do things with them.	4.36	0.931	5.53	0.910
6. I had the impression that I could be active in the environment of the DVR hotel English learning.	4.42	0.732	5.50	1.000
7. I felt like I could move around among the objects in the DVR hotel English learning environment.	4.47	0.845	5.22	0.832
8. It seemed to me that I could do whatever I wanted in the environment of the DVR hotel English learning.	3.89	0.523	5.28	0.815

Analysis of the descriptive statistics shows that both IVR and DVR improved students perceived spatial presence from the pre-test to the post-test. The DVR condition yielded higher post-test scores across nearly all items, with increases ranging from 0.5 to 1.4 points, compared to the more modest gains observed in the DVR condition (0.14–0.36 points). While some pre-test means were slightly higher for DVR, due to students' familiarity with desktop interfaces, the post-test results clearly indicate that IVR produced a stronger sense of spatial presence. This effect was particularly evident in items related to presence, interaction, and perceived control within the environment, suggesting that the immersive features of IVR more effectively enhanced students' spatial presence, likely through the experience of "being there" and the increased interactivity afforded by head-mounted VR. Although DVR provided a comfortable and familiar interface, its lower level of immersion limited the maximal spatial presence achievable during learning tasks.

2) Assumption testing

Before conducting the ANCOVA, the assumptions of normality, reliability, and homogeneity of variance were examined for the spatial presence scores. The Kolmogorov-Smirnov and Shapiro-Wilk tests indicated a slight deviation from normality ($p < 0.05$; see Table 5). However, given the relatively large sample size ($n = 144$), the distribution was considered approximately normal. This conclusion was further supported by visual inspection of the histogram (Fig. 3) and the Q-Q plot (Fig. 4). The spatial presence scale demonstrated acceptable internal consistency, with a Cronbach's alpha of 0.797. Levene's test confirmed that the assumption of homogeneity of variance was satisfied ($F = 0.166$, $p = 0.685$), indicating equal variance across the IVR and DVR groups.

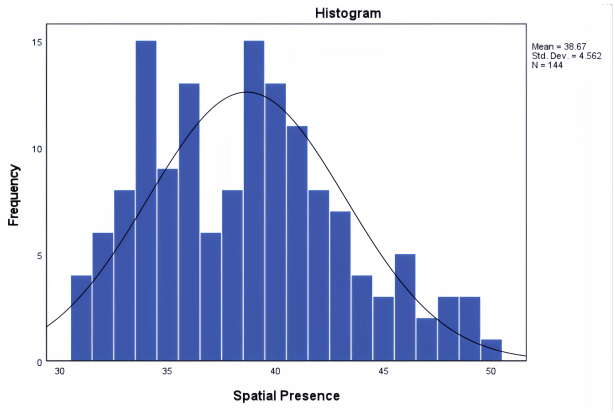


Fig. 3. Histogram of spatial presence scores.

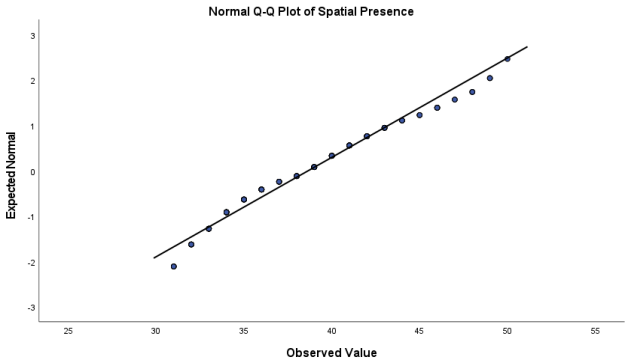


Fig. 4. Normal Q-Q plot of spatial presence scores.

3) ANCOVA results

An ANCOVA was conducted to examine the effect of the learning mode (IVR vs. DVR) on spatial presence, while controlling for the pre-test scores as a covariate. The results are presented in Table 6. The analysis indicated a significant effect of the learning mode on spatial presence, $F(1, 140) = 18.751$, $p < 0.001$, partial $\eta^2 = 0.118$, indicating that participants in the IVR group reported higher spatial presence than those in the DVR group.

Table 5. Tests of normality for spatial presence

Normality Test	Statistic	Df	Sig.
Kolmogorov-Smirnova	0.103	144	<0.001
Shapiro-Wilk	0.967	144	0.002

Table 6. ANCOVA results for spatial presence

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	744.965[a]	**3**	**248.322**	**17.634**	**<0.001**	**0.274**
Intercept	**239202.507**	**1**	**239202.507**	**16985.990**	**<0.001**	**0.992**
Mode	**264.063**	**1**	**264.063**	**18.751**	**<0.001**	**0.118**
Error	**1971.528**	**140**	**14.082**			
Total	**241919.000**	**144**				
Corrected Total	**2716.493**	**143**				

a: R Squared = 0.274 (Adjusted R Squared = 0.259).

B. Qualitative Findings

Semi-structured interviews were conducted with 8 participants to explore their perceived spatial presence during the IVR and DVR learning activities. Thematic analysis identified two core dimensions of spatial presence: Self-Location (SL) and Possible Actions (PA).

1) Self-Location (SL)

Students reported a strong sense of "being in" the virtual hotel environment, particularly under the IVR condition. One participant (R1) commented, "It felt like I was standing inside the hotel lobby instead of a classroom". Another student (R2) noted, "When I looked around, the environment felt natural, so I quickly forgot I was wearing a headset". Similarly, participant (R5) stated, "It felt like I was really facing a guest, not just practicing English". Additionally, participant (R8) mentioned, "The moment the scenario started, I felt like I had stepped into a real hotel setting, so my attention immediately shifted into the task". These accounts indicate that IVR created a more convincing experience of self-location compared with DVR.

2) Possible Actions (PA)

Students' perceptions of possible actions highlighted the immersive and interactive nature of IVR. Participants reported that IVR allowed them to freely interact with the virtual environment and feel a strong sense of agency. For example, participant (R3) commented, "I could interact with the virtual guest freely and try different phrases, which made the experience feel realistic and engaging". Participant (R4) noted, "I needed to think about my actions and the system responses simultaneously, but it made me more focused and involved in the task". Participant (R6) added, "The headset was slightly tiring after long sessions, but the immersive experience made it worth it". In contrast, participants reported that DVR was cognitively easier to control, allowing them to manage the conversation at their own pace. However, despite the relative comfort of DVR, IVR consistently provided a richer sense of possible actions, enabling higher engagement and perceived control during the tasks. These qualitative accounts align with the quantitative findings, confirming that IVR yielded higher post-test spatial presence than DVR.

V. Discussion

Overall, the findings support the primary hypothesis that IVR leads to higher levels of spatial presence than DVR in hotel English learning contexts. The integration of quantitative and qualitative findings provides a comprehensive understanding of students' spatial presence in the two VR learning environments. Quantitative results from the ANCOVA indicated that students in the IVR condition reported higher post-test spatial presence than those in the DVR condition, after controlling for pre-test scores. These results suggest that the immersive and interactive features of IVR, such as realistic 3D environments and enhanced sensory cues, may have facilitated stronger feelings of presence by supporting learners' spatial orientation and engagement during hotel English tasks. In contrast, the comparatively lower immersion of DVR may have constrained the sense of "being there," accounting for the lower spatial presence scores observed. Qualitative interviews corroborated these findings: students reported feeling more "located" in the virtual hotel environment and able to perform more realistic actions in IVR, whereas DVR users described the environment as less engaging and interactive.

These findings align with prior research emphasizing the benefits of IVR for enhancing spatial presence. For example, Wiepke [63] conducted a systematic review showing that IVR environments can improve learners' sense of presence and engagement across various tasks and disciplines. Similarly, Strand [64] reported that participants using head-mounted VR experienced significantly higher spatial presence than those in DVR environments, highlighting the role of immersion in creating a convincing sense of "being there". Furthermore, Lønne [65] further demonstrated that higher immersion in educational VR scenarios increased both spatial presence and positive affect, supporting the quantitative finding that IVR produced stronger spatial presence than DVR in the present study.

Qualitative interviews provided additional support for these findings. Students reported strong feelings of Self-Location (SL) in the IVR environment, describing the hotel scenario as realistic and attention- demanding. Regarding Possible Actions (PA), participants highlighted that IVR allowed for more interactive and realistic responses, enhancing their sense of agency and control within the environment. These observations are consistent with prior research. For example, Tang [66] noted that IVR can support engagement and perceived control, although higher immersion may also increase cognitive load. Similarly, Grecu [67] emphasized that while IVR can enhance interaction and learning, cognitive demands must be carefully managed to avoid fatigue or distraction.

From an educational perspective, these findings indicate that instructors should carefully balance immersion and cognitive load when integrating VR into hotel English learning. IVR appears particularly suitable for short, interactive exercises that require high engagement and strong spatial presence, while DVR may be more appropriate for extended practice or complex tasks where comfort and sustained attention are critical. Overall, the results extend previous VR instructional research by highlighting that immersive features enhance spatial presence and interactivity, but cognitive and usability considerations remain essential for effective learning [68].

VI. Conclusion and Recommendations

This study aims to investigate the impact of IVR and DVR on spatial presence in hotel English learning among Chinese university students. The findings indicate that, compared to DVR, IVR can generate a higher sense of spatial presence, highlighting the important role of immersive and interactive features in enhancing learners' feeling of "being there". Additionally, the qualitative results further corroborate and enrich the quantitative findings, revealing that IVR can also improve students' self-location and possible actions, thereby creating a richer and more engaging learning experience.

Compared to recent studies in VR-assisted language learning, this study is distinctive in situating IVR and DVR within a highly contextualized hotel English environment, thereby extending VR research into the domain of English for

Specific Purposes (ESP). Furthermore, the direct comparison of IVR and DVR while controlling for prior spatial presence, enhances the interpretability of immersion effects, which are rarely examined in language learning research. This approach positions the study at the intersection of immersive technology research, instructional design, and applied language pedagogy, offering both theoretical and practical insights.

A. Implications of the Study

This study has four main implications for the domain, technology, context, and target group.

Domain: The findings highlight how situational learning can be integrated with vocational application in hotel English courses, enhancing students' professional dialogue skills and workplace preparedness.

Technology: VR instructional tools can be strategically designed for different tasks: IVR most effective for short, interactive activities that maximize spatial presence, whereas DVR is better suited for longer, cognitively demanding tasks. These insights can guide educators and developers in optimizing VR design to improve learning effectiveness.

Context: IVR enhances spatial presence, interaction, and task engagement through realistic scenario simulation, while DVR supports extended practice and knowledge consolidation. Aligning the level of immersion with task type and cognitive demands can optimize learning outcomes.

Target group: First-year hotel management students with limited VR experience benefit from tailored VR-based ESP instruction. IVR fosters spatial presence and interactive dialogue skills, whereas DVR provides a stable platform for reinforcing language knowledge. Instructional design should consider learners' prior experience and cognitive load to maximize instructional effectiveness.

B. Limitations and Future Work

Despite the meaningful findings of this study, several limitations should be acknowledged. First, the observed advantages of IVR may be partially influenced by a novelty effect, as most participants had limited prior exposure to VR technology. This heightened situational interest may temporarily inflate learners perceived spatial presence in the short term and does not necessarily guarantee sustained engagement or long-term learning benefits. Second, the study was conducted within a relatively short instructional period, which constrained the examination of whether spatial presence effects persist over time or translate into durable language performance and professional skill transfer. Third, the sample was drawn from a single academic program at one institution, which restricts the generalizability of the findings to other learner populations, educational contexts, and levels of prior VR experience.

To address these limitations, future research should adopt a multi-pronged approach. Longitudinal or cross-sectional studies comparing VR-naïve and VR-experienced learners are needed to disentangle the temporary appeal of novelty from the genuine pedagogical efficacy of immersive technology. Design-based research can be conducted to develop and test instructional models that are intrinsically engaging within VR environments. Moreover, future studies should include larger and more heterogeneous participant groups, encompassing students from different academic years, institutions, cultural backgrounds, and varying levels of prior VR experience. Finally, longitudinal tracking over a semester or academic year is recommended to assess not only the retention of spatial presence but also the sustained development of professional dialogue skills and their transfer to simulated or real workplace scenarios, thereby evaluating the ultimate vocational impact of VR training.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Meng Huan conceptualized the study, developed the methodology, conducted the research, collected and analyzed the data, drafted the manuscript, and prepared the tables and figures; Mageswaran Sanmugam contributed to the theoretical framework and methodology, provided academic guidance, and critically reviewed and edited the manuscript. All authors approved the final version of the manuscript.

References

[1] T. P. Wu, H. C. Wu, G. Ye, X. Wu, and B. Pan, "The contribution of international tourism development to economic growth in Chinese economy," *Journal of Policy Research in Tourism, Leisure and Events*, vol. 16, no. 4, pp. 575–589, 2024. doi: 10.1080/19407963.2022.2088773

[2] S. Prombut, W. Winaitham, N. Prakongchati, K. Phuwichit, T. Klomkaew, and P. Niyomkarn, "The requirement of English communication skills of the hotel receptionists at Kamphaeng Phet Province," *The Journal of Sirindhornparithat*, vol. 25, no. 1, 2024.

[3] I. Ma'fiyah and S. Sumardiono, "English language needs for hotel service and ESP program: The case of hospitality vocational secondary school," *Journal of Innovation in Educational and Cultural Research*, vol. 4, no. 3, pp. 429–437, 2023. doi: 10.46843/jiecr.v4i3.680

[4] X. Pan and W. Zhong, "Teaching spoken English for hospitality management major in China," *Journal of Hospitality, Leisure, Sport & Tourism Education*, vol. 30, 100375, 2022. doi: 10.1016/j.jhlste.2022.100375

[5] Z. Xu, S. Jain, and M. Kankanhalli, "Hallucination is inevitable: An innate limitation of large language models," arXiv Preprint, arXiv:2401.11817, 2024. doi: 10.48550/arXiv.2401.11817

[6] F. N. Fazal, M. S. Tahir, M. Chaudhary, and M. Abbasi, "Effectiveness of AI integration into Computer Assisted Language Learning (CALL) on student writing skills based on gender," *Pakistan Journal of Humanities and Social Sciences*, vol. 12, no. 1, pp. 224–230, 2024.

[7] H. R. Hoffman, "The impact of immersive VR versus desktop VR technologies on learning attainment and sense of presence," M.S. thesis, Univ. of South Dakota, Vermillion, SD, 2023.

[8] Z. Wei, J. Liao, L. H. Lee, H. Qu, and X. Xu, "Towards enhanced learning through presence: A systematic review of presence in virtual reality across tasks and disciplines," arXiv Preprint, arXiv:2504.13845, 2025. doi: 10.48550/arXiv.2504.13845

[9] F. Çelik and M. H. Baturay, "The effect of metaverse on L2 vocabulary learning, retention, student engagement, presence, and community feeling," *BMC Psychology*, vol. 12, 58, 2024. doi: 10.1186/s40359-024-01549-4

[10] S. Alhawsawi and F. Alzaid, "Systematic review of virtual reality's uses on vocabulary learning in higher education," *Smart Learning Environments*, vol. 12, 60, 2025. doi: 10.1186/s40561-025-00414-0

[11] N. Owie and M. Alizadeh, "Virtual reality for language learning," *ELT Journal*, pp. 397–400, 2025. doi: 10.1093/elt/ccaf025

[12] X. Hu and S. Wang, "The effects of VR integration in teaching English as a foreign language: Academic achievement, motivation, and engagement among middle school learners," *International Journal of Research and Innovation in Social Science*, vol. 8, pp. 1833–1841, 2025. doi: 10.47772/IJRISS.2025.908000153

[13] M. K. Thakur and A. Bhardwaj, "The role and importance of effective English communication skills in tourism & hotel industries," *Journal of the Asiatic Society of Mumbai*, vol. 46, no. 3, pp. 97–101, 2023. doi: 10.13140/RG.2.2.18154.49608

[14] M. S. Husin, R. Ibrahim, G. de Mello, and K. Ariffin, "Perceptions on English communication skills training for hotel front office

employees," *European Proceedings of Educational Sciences*, pp. 773–782, 2023. doi: 10.15405/epes.23097.69
[15] M. Yahya, K. A. Ekasani, R. Ratnah, N. S. Wahyuningsih, F. Saputri, S. Supriatnoko, and Y. D. Fitriansyah, *English for Hospitality*, Jakarta, Indonesia: PT. Green Pustaka Indonesia, 2025.
[16] Y. T. Guntoro, "ESP course design for the Indonesian hotel industry: Need analysis," *Englie: English Learning Innovation*, vol. 2, no. 2, pp. 52–61, 2021. doi: 10.22219/englie.v2i2.17041
[17] A. H. Miqawati, F. Wijayanti, I. H. Al Mujib, A. Kadhafi, and I. Devy, "Identifying English communication competence for front office of hotel industries," in *Proc. ICoSHIP 2023: Proceedings of the 4th International Conference on Social Science, Humanity and Public Health*, Surabaya, East Java, Indonesia, 18–19 Nov. 2023.
[18] D. Evangelou, M. Mulders, and B. Sekerci, "An overview of commercial virtual reality providers in education: Mapping the current market landscape," *Applied Sciences*, vol. 15, no. 7, 3906, 2025. doi: 10.3390/app15073906
[19] C. N. Akpen, S. Asaolu, S. Atobatele, H. Okagbue, and S. Sampson, "Impact of online learning on student's performance and engagement: A systematic review," *Discover Education*, vol. 3, 205, 2024. doi: 10.1007/s44217-024-00253-0
[20] M. B. Purwanto, R. Hartono, and S. Wahyuni, "Essential skills challenges for the 21st century graduates: Creating a generation of high-level competence in the Industrial Revolution 4.0 era," *Asian Journal of Applied Education*, vol. 2, no. 3, pp. 279–292, 2023. doi: 10.55927/ajae.v2i3.3972
[21] P. Lertchalermtipakoon, U. Wongsubun, and P. Kawinkoonlasate, "Need analysis: English language use by tourism and hospitality industry students," *English Language Teaching*, vol. 14, no. 3, pp. 59–68, 2021. doi: 10.5539/elt.v14n3p59
[22] M. Sun, X. Pan, K. Yuan, M. P. Vortia, and Y. Sun, "A study on blended teaching models based on symbiosis theory: Taking shipping-related course instruction as an example," *Journal of Educational Theory and Practice*, vol. 3, no. 1, 2026. doi: 10.62177/jetp.v3i1.1024
[23] N. Chhabhaiya, B. Patle, and P. Bhojane, "Virtual and augmented reality applications: A broader perspective review," *Journal of Data Science and Intelligent Systems*, vol. 4, no. 1, pp. 13–25, 2026. doi: 10.47852/bonviewJDSIS32021064
[24] M. K. Noordin and M. A. Azmi, "Exploring the potential benefits and challenges of augmented reality in TVET programs: A mixed methods study," *International Journal for Research and Innovation in Social Science*, pp. 7363–7373, 2025. doi: 10.47772/IJRISS.2025.909000601
[25] M. R. Jalilvand and H. Ghasemi, "Augmented reality technology in tourism and hospitality research: A review from 2010 to 2024," *Journal of Science and Technology Policy*, 2024.
[26] M. I. Baig and E. Yadegaridehkordi, "Flipped classroom in higher education: A systematic literature review and research challenges," *International Journal of Educational Technology in Higher Education*, vol. 20, 61, 2023. doi: 10.1186/s41239-023-00430-5
[27] S. O. Thorp, L. M. Rimol, S. Lervik, H. R. Evensmoen, and S. Grassini, "Comparative analysis of spatial ability in immersive and non immersive virtual reality: The role of sense of presence, simulation sickness, and cognitive load," *Frontiers in Virtual Reality*, vol. 5, 1343872, 2024. doi: 10.3389/frvir.2024.1343872
[28] P. D. Acevedo, A. J. Magana, B. Benes, and C. Mousas, "A systematic review of immersive virtual reality in STEM education: Advantages and disadvantages on learning and user experience," *IEEE Access*, vol. 12, pp. 189359–189386, 2024.
[29] G. Makransky and R. E. Mayer, "Benefits of taking a virtual field trip in immersive virtual reality: Evidence for the immersion principle in multimedia learning," *Educational Psychology Review*, vol. 34, no. 3, pp. 1771–1798, 2022. doi: 10.1007/s10648-022-09675-4
[30] A. Sudár and Á. B. Csapó, "Comparing desktop 3D virtual reality with web 2.0 interfaces: Identifying key factors behind enhanced user capabilities," *Heliyon*, vol. 10, e31717, 2024. doi: 10.1016/j.heliyon.2024.e31717
[31] R. Kaplan Rakowski, L. Lin, and T. Wojdynski, "Learning vocabulary using 2D pictures is more effective than using immersive 3D stereoscopic pictures," *International Journal of Human-Computer Interaction*, vol. 38, no. 4, pp. 299–308, 2022. doi: 10.1080/10447318.2021.1938394
[32] H. C. Lum, J. R. Lee, S. Evans, Y. Liu, and T. Kampf, "Impact of immersion in Virtual Reality (VR) and desktop environments on spatial ability," in *Proc. Human Factors and Ergonomics Society Annual Meeting*, vol. 68, 2024, pp. 1577–1581. doi: 10.1177/10711813241265656
[33] C. Castro, A. I. Dinu, P. Doncel, M. B. Sánchez-Barrera, M. Á. Muñoz, F. M. Gómez-Campos, and M. A. Trógolo, "Immersive tendencies and presence in virtual reality: Spanish adaptation of the ITQ and IPQ, and their relationship with suggestibility and cybersickness," *Virtual Reality*, vol. 30, no. 1, 49, 2026. doi: 10.1007/s10055-025-01301-2
[34] R. C. A. Barret, R. Poe, J. W. O'Camb, C. Woodruff, S. M. Harrison, K. Dolguikh, and M. R. Blair, "Comparing virtual reality, desktop-based 3D, and 2D versions of a category learning experiment," *PLOS One*, vol. 17, no. 10, e0275119, 2022. doi: 10.1371/journal.pone.0275119
[35] E. Chauvergne, M. Hachet, and A. Prouzeau, "User onboarding in virtual reality: An investigation of current practices," in *Proc. 2023 CHI Conf. Hum. Factors Comput. Syst.*, 2023, pp. 1–15. doi: 10.1145/3544548.3581211
[36] J. Oh, J. Kang, and Y. Kim, "The effectiveness of virtual reality with or without head-mounted displays: Focusing on the role of spatial presence and immersive tendency for responding to particulate matter," *Virtual Reality*, vol. 29, no. 2, 80, 2025. doi: 10.1007/s10055-025-01156-7
[37] S. Triberti, C. Sapone, and G. Riva, "Being there but where? Sense of presence theory for virtual reality applications," *Humanities and Social Sciences Communications*, vol. 12, no. 1, pp. 1–6, 2025. doi: 10.1057/s41599-025-04380-3
[38] J. M. Juliano, N. Schweighofer, and S. Liew, "Increased cognitive load in immersive virtual reality during visuomotor adaptation is associated with decreased long-term retention and context transfer," *Journal of NeuroEngineering and Rehabilitation*, vol. 19, 106, 2022. doi: 10.1186/s12984-022-01084-6
[39] R. G. Pramasta and Y. W. A. Wibowo, "Virtual reality and language learning in hospitality: An investigation of students' perceptions," *International Journal of Current Science Research and Review*, vol. 6, no. 8, pp. 5647–5652, 2023. doi: 10.47191/ijcsrr/V6-i8-32
[40] I. Jung, "Integrating VR tours in online language learning: A design-based research study," *Australasian Journal of Educational Technology*, vol. 41, no. 1, pp. 60–87, 2025. doi: 10.14742/ajet.9498
[41] W. Yan, V. L. Lowell, and L. Yang, "Developing English language learners' speaking skills through applying a situated learning approach in VR-enhanced learning experiences," *Virtual Reality*, vol. 28, no. 4, 167, 2024. doi: 10.1007/s10055-024-01061-5
[42] A. Degenhard, A. Askari, M. Rietzler, and E. Rukzio, "When do we feel present in a virtual reality? Towards sensitivity and user acceptance of presence questionnaires," in *Proc. of the 2025 CHI Conference on Human Factors in Computing Systems,* 2025, pp. 1–29.
[43] B. Yan, S. Ni, X. Wang, J. Liu, Q. Zhang, and K. Peng, "Using virtual reality to validate the Chinese version of the independent television commission sense of presence inventory," *SAGE Open*, vol. 10, no. 2, 2158244020922878, 2020. doi: 10.1177/2158244020922878
[44] M. Lombard, T. B. Ditton, and L. Weinstein, "Measuring presence: The Temple Presence Inventory (TPI)," in *Proc. Int. Soc. Presence Research Annual Conf. (ISPR)*, vol. 15, pp. 1–15, 2009.
[45] T. Hartmann, W. Wirth, and H. Schramm *et al.*, "The Spatial Presence Experience Scale (SPES): A short self report measure for diverse media settings," *Journal of Media Psychology*, vol. 28, no. 1, pp. 1–15, 2016. doi: 10.1027/1864-1105/a000137
[46] W. Wirth, T. Hartmann, S. Boecking, P. Vorderer, P. Klimmt, H. Schramm *et al.*, "A process model of the formation of spatial presence experiences," *Media Psychology*, vol. 9, pp. 493–525, 2007. doi: 10.1080/15213260701283079
[47] S. M. Alessi and S. R. Trollip, *Multimedia for Learning: Methods and Development*, 3rd ed. Boston, MA, USA: Allyn & Bacon, 2001.
[48] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence: Teleoperators and Virtual Environments*, vol. 7, no. 3, pp. 225–312, 1998. doi: 10.1162/105474698565686
[49] T. W. Schubert, "A new conception of spatial presence: Once again, with feeling," *Communication Theory*, vol. 19, no. 2, pp. 161–187, 2009. doi: 10.1111/j.1468-2885.2009.01340.x
[50] M. H. Draper, D. B. Kaber, and J. M. Usher, "Telepresence," *Human Factors in Virtual Environments (Technical Report)*, 1998. doi: 10.1518/001872098779591386
[51] J. Steuer, "Defining virtual reality: Dimensions determining telepresence," *Journal of Communication*, vol. 42, no. 4, pp. 73–93, 1992.
[52] M. Slater and A. Steed, "A virtual presence counter," *Presence:* Teleoperators *& Virtual Environments*, vol. 9, no. 5, pp. 413–434, 2000. doi: 10.1162/105474600566925
[53] T. W. Schubert, F. Friedmann, and H. Regenbrecht, "The experience of presence: Factor analytic insights," *Presence: Teleoperators and Virtual Environments*, vol. 10, pp. 266–281, 2001.
[54] M. C. C. S. de Carvalho *et al.*, "Assessment of spatial presence experiences and symptoms of cybersickness in dynamic and static virtual environments," *Journal of Clinical Medicine*, vol. 14, no. 12, 4200, 2025. doi: 10.3390/jcm14124200

[55] H. Rokhaniyah, "Developing web-based online test system to boost IELTS academic reading score," *English Review: Journal of English Education*, vol. 9, no. 2, pp. 9–18, 2021.
[56] N. V. Ivankova, J. W. Creswell, and S. L. Stick, "Using mixed-methods sequential explanatory design: From theory to practice," *Field Methods*, vol. 18, no. 1, pp. 3–20, 2006.
[57] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Lawrence Erlbaum Associates, 1988. doi: 10.4324/9780203771587
[58] H. Hasan, A. Tumiran, and M. A. Tumiran, "Determining adequate sample size for social survey research," *Quantum Journal of Social Sciences and Humanities*, vol. 5, no. 2, pp. 70–86, 2024. doi: 10.55197/qjssh.v5i2.346
[59] A. Field, *Discovering Statistics Using IBM SPSS Statistics*, 5th ed. Sage, 2018.
[60] B. G. Tabachnick and L. S. Fidell, *Using Multivariate Statistics*, 7th ed. Pearson, 2019.
[61] G. Guest, A. Bunce, and L. Johnson, "How many interviews are enough? An experiment with data saturation and variability," *Field Methods*, vol. 18, no. 1, pp. 59–82, 2006. doi: 10.1177/1525822X05279903
[62] C. B. Califf, S. Sarker, and S. Sarker, "The bright and dark sides of technostress: A mixed methods study involving healthcare IT," *MIS Quarterly*, vol. 44, no. 2, pp. 809–856, 2020. doi: 10.25300/MISQ/2020/14818
[63] A. Wiepke and B. Heinemann, "A systematic literature review on user factors to support the sense of presence in virtual reality learning environments," *Computers & Education: X Reality*, vol. 4, 100064, 2024. doi: 10.1016/j.cexr.2024.100064
[64] J. R. Strand, D. D. Huang, C. M. Clark, Y. Shahrbanian, and J. T. Hansen, "The effect of immersion on sense of presence and affect when experiencing an educational scenario in virtual reality: A randomized controlled study," *PLOS One*, vol. 18, no. 11, e10285157, Nov. 2023.
[65] T. F. Lønne, H. R. Karlsen, E. Langvik, and I. S. Lehouillier, "The effect of immersion on sense of presence and affect when experiencing an educational scenario in virtual reality: A randomized controlled study," *Heliyon*, vol. 9, e17196, 2023. doi: 10.1016/j.heliyon.2023.e17196
[66] Z. Tang, Y. Wang, H. Liu, Q. Liu, and S. Jiang, "Experiencing an art education program through immersive virtual reality or iPad: Examining the mediating effects of sense of presence and extraneous cognitive load on enjoyment, attention, and retention," *Frontiers in Psychology*, vol. 13, 957037, 2022. doi: 10.3389/fpsyg.2022.957037
[67] V. Grecu *et al.*, "The impact of immersive virtual reality instruction on cognitive demands in novice learners," *Applied Sciences*, vol. 15, no. 23, 12534, 2025. doi: 10.3390/app152312534
[68] H. C. Lum, J. R. Lee, S. Evans, and Y. Liu, "Impact of immersion in Virtual Reality (VR) and desktop environments on spatial ability," in *Proc. of the Human Factors and Ergonomics Society Annual Meeting*, vol. 68, no. 1, 2024, pp. 1577–1581. doi: 10.1177/10711813241265656