



A Research on the Building of Virtual Scene Situations and the Promotion of Cross-Culture Communicative Abilities in Immersion-type English Teaching

Juan Yang¹ and Anyi Zhou^{2,*}

1 School of Foreign Studies, Xiangsihu College of Guangxi Minzu University, Nanning, 530225, Guangxi, China

2 School of Asian Studies, Guangxi University of Finance and Economics, Nanning, 530007, Guangxi, China

SUMMARY: *For the problems of lacking situation truth, few culture hints, and not strong cross-culture shifting in university English teaching rooms, this research builds a virtual reality based immersion English teaching frame which contains culture hints, branch conversations, partner together existing, and reflection feed back. One 12-week quasi-experimental research was carried out at one university in Guangxi, China. This research has 126 first-year undergraduate students who study majors that are not English disciplines. We have carried out the evaluation through the combination of Intercultural Communication Competence (ICC) Scale, oral task score, behavior record, Immersion and Cognitive Load Scale, and 4-week post-test. The outcomes have indicated that the ICC total score of the experimental group has a rise from 3.21 to 4.06, with the post-test score being maintained at 3.91; as for the control group, its post-test examination score is 3.58. The experimental group's oral task scores had an increase from 71.8 to 85.6, average reaction time was decreased to 1.65 s, and correction success rate had a rise to 77.3%. The results which belong to specific scenarios show that the highest transfer retention rates have been observed in the scenarios of project negotiation and host family dinner; hence academic seminar Q&A is more significantly influenced by the load of conceptual retrieval. The cutting experiment further showed that the cultural hint layer and branch conversations have the most prominent contribution to cross-cultural suitability, interaction density and conversion efficiency. This research proves that the effect of VR in English immersion teaching is dependent on the cooperative arrangement of high-quality scene building, interaction design, and cognitive load management, and hence offers a repeatable assessment route for the change of immersion teaching and the building of scene banks in college English classes.*

KEYWORDS: *virtual reality; English immersive teaching; scenario construction; intercultural communicative competence; higher education*

1 Introduction

As university English instruction has entered a new phase that emphasizes both application-oriented and competency-oriented approaches, the focus of classroom assessment has gradually shifted from vocabulary memorization, grammatical accuracy, and text retelling to sustained expression, pragmatic adjustment, and cross-cultural negotiation skills in complex scenarios.

*zhoukeith0123@126.com

<https://doi.org/10.65102/is2026756>

For most non-English majors, the challenge in conventional classrooms is not merely an inability to "speak," but rather a difficulty in maintaining consistent output during tasks that approximate real-life interactions: when placed in specific scenarios such as reception interactions, classroom discussions, public service procedures, or project negotiations, students often manage to convey basic information but struggle to adjust their expression in a timely manner based on the interlocutor's identity, relational distance, topic sensitivity, and cultural expectations. When communication shifts from the application of linguistic knowledge to the articulation of positions, politeness negotiation, resolution of misunderstandings, and value judgments, learners who previously performed adequately in written exercises frequently encounter issues such as response stagnation, limited strategic flexibility, pragmatic imbalance, and insufficient cultural interpretation. It is exactly on this account that cross-cultural communication abilities are more and more regarded as a core constituent of real language ability in college English teaching rooms, not only extra content connected to language cultivation [1, 2].

This pedagogical pressure stems primarily from changes in the application scenarios themselves. The students currently enrolled in university English courses are increasingly participating in online international courses, joint projects, exchange programs, overseas internships, and cross-cultural team collaboration. In these tasks, students face not single-round Q&A exchanges, but dynamic judgment within continuous interaction: when to explain context, when to soften tone, when to confirm the other party's intent through follow-up questions, and when to express disagreement without damaging the relationship. Although traditional classrooms can supplement some contextual elements through text materials, video case studies, and role-playing, such training often remains at the level of "knowing what to do." Learners' insufficient perception of spatial cues, recipient feedback, emotional pressure, and multi-turn interactions means that correct expressions in the classroom may not necessarily transfer to real-life communication. Addressing this gap, the question of how to construct a language environment in teaching that more closely resembles real-life interaction-and how to integrate cultural judgment with linguistic action within a single task unit-has become an unavoidable issue in English language teaching research.

Virtual reality (VR) technology, which has immersive space environment, real-time interaction function and can program scenes, therefore provides a new technical method for solving this difficult problem. When we compare VR with traditional video materials, VR changes learners' role from that of onlookers to that of joiners, therefore it allows them to finish language work inside an environment where they can feel, give reaction to, and make adjustment to their own behaviors. Currently existing researches show that the function of VR in foreign language studying is not any more restricted to "boosting interest" or "strengthening novelty." Its more important meaning lies in its capability to arrange scenes, behaviors, object responses, and nonverbal signals into a continuous mutual environment, thus making language production no longer take place separately from particular situations [4-8]. This change is especially vital for spoken practice, because language expression in actual daily communication is never finished alone without the talking object and cultural setting background. Learners who study must find out problems which are caused by the situation, arrange their words as the interactive process goes on, hence quickly change their methods after they get responses. In other terms, VR is not only give a visual replacement but also an environment technology which has ability to rebuild communication conditions of classroom.

Direct investigations upon cross-cultural learning are moreover going forward side by side. In the past few years, correlative researches have already started to utilize VR or social VR in cross-cultural curricula, overseas students' adaptation, cultural teaching, and interactive participation researches, with a focus put on dimensions including cultural consciousness,

openness, interactive participation, negotiation abilities, and study participation [9-11]. These results prove that VR has already built a firm base for getting into the main current of cross-cultural communication teaching. In the meanwhile, studies within the domain of language study have additionally uncovered the influence of VR upon speaking nervousness, communicative initiative, cognitive engrossment, and immersed experiences [12-16]. Put together, these researches indicate a more concrete problem: when learners are put into virtual situations that have culture limits and topic feedback, language teaching in the class starts to show behavior models that are different from those of traditional teaching methods. Learning results already not only depend on the whole completeness of knowledge passing, but more and more hinge on whether scenario design, interaction density and feedback structures are appropriate.

However, existing research still has several shortcomings that directly affect the focus of this paper. First, while many studies report "immersion" as a general advantage, few go on to investigate which specific instructional components contribute to these benefits. In the context of cross-cultural English classrooms, the key variables influencing learning outcomes are not merely whether the technology is immersive, but also whether cultural cues are sufficient, whether dialogues feature branching paths, whether peers are actively engaged, and whether teacher feedback is integrated into post-scenario reflection. If these factors are not disentangled, the educational value of VR classrooms is likely to remain at the level of general impressions, making it difficult to translate into reusable teaching modules. Second, some studies focus more on oral fluency, reduced anxiety, or learning acceptance, while insufficiently discussing "the specific pathways through which intercultural communication skills are enhanced." The fact that students speak more in virtual environments does not necessarily imply that their abilities in cultural interpretation, relationship management, and strategy adjustment have improved in tandem. Furthermore, the contextual applicability of existing research remains inconsistent. Many experiments rely on one-time experiences or short-term tasks, with selected scenarios primarily designed for demonstration purposes; few are designed as reusable classroom units. Consequently, while research conclusions may illustrate the technology's potential, they struggle to directly address how to integrate it into university English courses over the long term, how to evaluate its effectiveness, how to compare outcomes, and how to implement it. Finally, immersive teaching is accompanied by practical constraints such as cognitive load management, device adaptation, the waning novelty effect, and the organizational costs for instructors. If research merely reports post-test improvements without discussing these boundary conditions, it is difficult to arrive at truly actionable conclusions for curriculum design.

It is against this backdrop that this paper narrows the research question to: How can we construct a VR scenario unit in university English immersive instruction that can consistently support cross-cultural interaction tasks, enabling language expression training, cultural cue recognition, strategy adjustment, and reflective feedback to occur within the same classroom cycle, and ultimately translate into measurable improvements in cross-cultural communication competence? Addressing this issue, this paper does not treat VR as an external presentation tool but rather as a mechanism for organizing situational contexts within the English classroom. The core of the research does not lie in proving the broad proposition of "whether the use of VR is effective," but rather in explaining: what kind of situational construction is more suitable for cross-cultural English teaching; which instructional components truly generate gains; in which competency dimensions these gains manifest; and whether such a design possesses stability in curricular implementation.

Based on this, this study undertakes the following tasks. First, focusing on high-frequency communicative tasks in university English classrooms, we construct a VR-based immersive teaching framework comprising cultural cues, branching dialogues, peer co-presence, teacher

scaffolding, and reflective feedback, and organize it into reusable contextualized task units. Second, in quasi-experimental teaching, we integrate assessment scales, oral tasks, behavioral logs, and interview data to conduct a comprehensive evaluation of changes in learners' cross-cultural communication competence, interactive behavior, and situational experience. Third, through scenario-specific comparisons and modular ablation analysis, we identify the key contextual components influencing learning outcomes and further discuss the application boundaries and implementation value of this framework in university English classrooms. Through the above work, this paper aims to provide a research design for the integration of virtual reality technology into immersive English instruction that is more closely aligned with actual course settings, features a closed-loop evaluation system, and is replicable.

2 Methods

2.1 Research Context, Participants, and Multi-source Data

This research was conducted in a general English class at a university in Guangxi, China, and it used a whole-class quasi-experiment design. The starting sample was made up of 132 first-year undergraduate students who study majors which are not English. After we exclude those participants whose absentee rate is more than 20% or who have no complete measurement data, the final sample includes 126 students, which are divided into one experimental group ($n = 63$) and one control group ($n = 63$). These two groups all obtained teaching from the identical instruction group. The course subjects, all teaching hours, teaching contents, and evaluation items were kept unchanged; The only difference existed in the carry out way of the learning situations: the experiment group used immersion teaching constructed on VR situations, while the contrast group used traditional situation-based teaching that includes video introductions, PowerPoint situation hints, and face-to-face role playing. The duration of our study was 12 weeks, and a follow-up assessment was conducted by us 4 weeks after the intervention had been concluded. The procedures of informed consent and anonymization were finished before the semester began. The features of the research participants and baseline equality examinations are displayed in Table 1.

Table 1: Participant Individual Features and Baseline Equality Checking Tests

Variable	Experimental Group (n=63)	Control Group (n=63)	Statistic Value	<i>p</i>
Age (years)	19.42±0.68	19.37±0.71	t=0.41	0.684
Female Proportion (%)	54.0	52.4	$\chi^2=0.03$	0.857
Entrance English Level Score	72.4±8.1	71.8±7.9	t=0.43	0.671
Pre-Test ICC Total Score (1-5)	3.21±0.38	3.19±0.41	t=0.28	0.782
Pre-Test Speaking Task Score (0-100)	71.8±7.4	72.1±7.1	t=-0.25	0.804
Previous VR Usage Experience (%)	23.8	25.4	$\chi^2=0.04$	0.841

Table 1 displays that there were not any marked differences between the two groups on the aspects of age, English ability, pre-test ICC marks, pre-test oral marks, and past VR use experience, hence it establishes the baseline conditions which are needed for later comparisons.

The sample data consisted of five sources. First, the Intercultural Communication Competence Scale was used to track changes in students' open-mindedness, cultural knowledge and interpretation, interaction skills, strategic adaptation, and critical cultural awareness. Second, the Spoken Task Scoring Sheet was used to evaluate students' semantic coherence, pragmatic appropriateness, negotiation skills, and corrective abilities in specific contexts. Third,

a system behavior log was used to record task completion time, response latency, turn density, frequency of cultural cue invocation, and correction events. Fourth, the Presence and Cognitive Load Scale was used to characterize the immersive experience and its associated costs. Fifth, learning reflection forms and semi-structured interviews were used to elucidate the behavioral mechanisms underlying the numerical results. The mapping relationships among the collection, coding, and evaluation of multi-source data are shown in Figure 1.

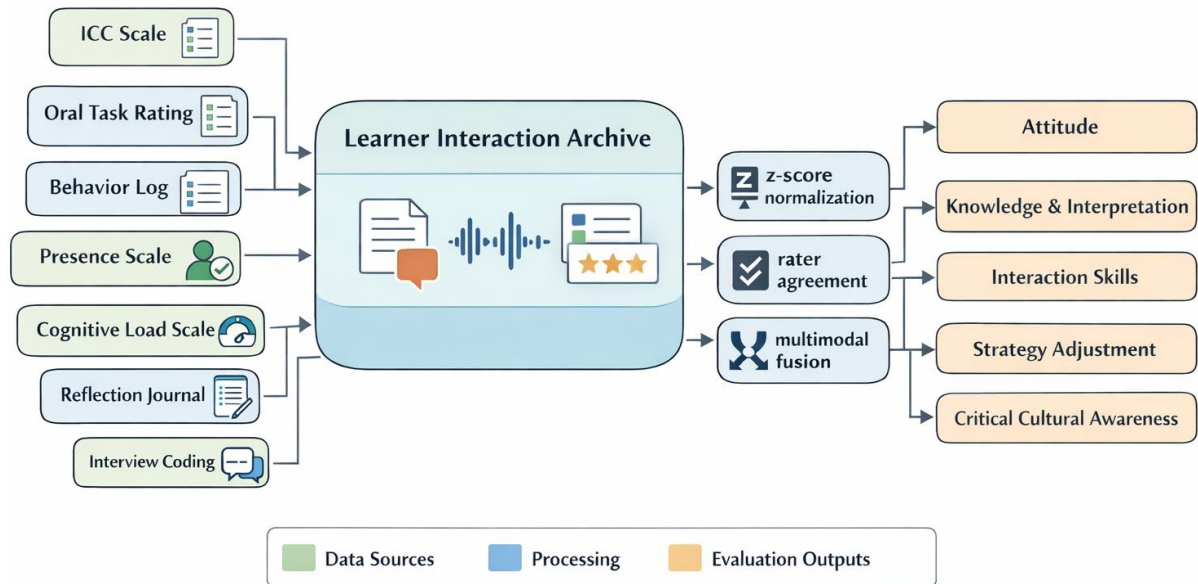


Figure 1: Mechanism for multi-source data collection, coding, and evaluation mapping.

For the promotion of data dependability, the spoken language tasks have utilized double-blind marking from two raters. Before the scoring work, we used 15 samples to finish the calibration of standards, therefore the inter-rater reliability (ICC) which got after the formal scoring was 0.87. The calling of cultural cues, correction methods, and turn-taking structures in the behavior records were separately coded by two researchers; We extracted a 20% sample to carry out consistency check, therefore it got a Cohen's κ value of 0.86. Interview record materials were subjected to thematic coding work, with the focus placed on four major axes: "situation-level involvement," "culture-based judgment," "expression-related revision," and "technology-related burden."

2.2 VR Scenario Design and Immersive Instructional Intervention

In line with the research questions of this study, the VR scenario units were organized into three hierarchical levels: "Scenario Framework-Interaction Trigger-Reflection and Transfer." The scenario-carrying layer provides spatial cues, character identities, and action tasks; the interaction-triggering layer generates dialogue branches, situational conflicts, and feedback nodes; and the reflection-transfer layer makes language choices and cultural judgments from the current interaction explicit and incorporates them into subsequent tasks. To illustrate how scenario construction is coupled with language tasks and cultural cues, the intervention mechanism of this study is shown in Figure 2.

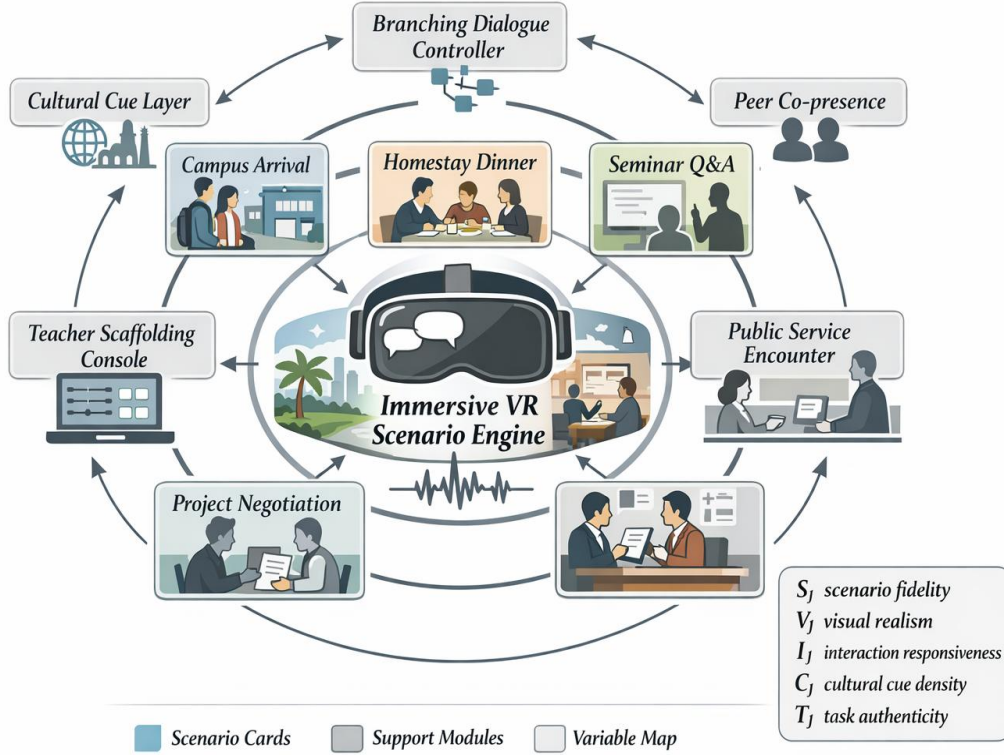


Figure 2: Context Construction and Task Mapping Mechanism in VR English Immersive Instruction.

The quality of scenario construction is measured by the scenario fidelity index S_j , as shown in Equation (1).

$$S_j = \alpha V_j + \beta I_j + \gamma C_j + \delta T_j \quad (1a)$$

$$\alpha + \beta + \gamma + \delta = 1 \quad (1b)$$

In the equation, S_j represents the comprehensive fidelity of the j th scenario unit j ; V_j denotes visual-spatial fidelity; I_j indicates interactive responsiveness; C_j represents cultural cue density; T_j denotes task authenticity; α , β , γ , and δ are weights. Based on the evaluations of five experts in English language teaching and educational technology, this study adopted $\alpha = 0.25$, $\beta = 0.25$, $\gamma = 0.30$, and $\delta = 0.20$. The mean S_j score for the five scenario units was 4.18, with "Project Negotiation" scoring the highest (4.31) and "Academic Seminar Q&A" the lowest (4.05); these differences primarily stemmed from interaction pressure and the complexity of cultural judgments.

From the perspective of specific design, this research has set up five core situations: orientation work for international students, having dinner with a host family, question and answer in an academic seminar, public service flow processes, and cross-cultural project business negotiations. Each scene inside contains one main work item and two extra non-required small tasks. Regarding the example of "Dinner with a Host Family", the main task is that people must carry out table conversation and give polite replies, while the sub-tasks include the explanation of food taboos and the negotiation of schedules; This system gives cultural hints to students by character talks, things in environment and event activations, for example addressing choices, refusing methods, tone adjusting words and time arrangement of non-language answers. Taking "item negotiation" as one example, the system puts key point on

strengthening position statement, dealing with fund disagreement, and compromise methods, it requires students to clarify their standpoints, harmonize benefits, and reach common decisions via many turns of talk.

This interference includes one 100-minute meeting every week, in all for 12 weeks. Week 1 is given to baseline measure and equipment adaptation; In the 2nd to 11th weeks, rotating teaching is carried out among five kinds of situation types; and Week 12 brings the study to an end through post-tests and interviews. For handling the novelty influence and early cognition burden that come with immersive apparatuses, the difficulty level of tasks was on purpose lowered in the first two weeks. At first, students got themselves familiar with the operation of headset, the switching of scenes, and the basic feedback mechanisms; after that, the multi-role interactions and cultural conflict tasks were gradually introduced to them [17]. In the experiment group, each class segment goes along in this sequence: "teacher explanation-VR mutual operation-classmate discussion conclusion-teacher direction-thinking and writing"; the control group, however, used video examples, text materials and offline role-playing to learn the same contents. The newly done VR study which bases on scenarios shows that task true degree, professional situation insertion, and instant response can greatly raise task participation degree and the quality of language shifting [18-20]. Based on this point, our this study put its scenario design focus on "communicative pressure" and not on "spectator-oriented scenarios."

2.3 Measurement Tools, Statistical Models, and Evaluation Protocol

The Cross-Cultural Communication Competence Scale consists of 25 items rated on a 5-point Likert scale, organized into five dimensions: open-mindedness, cultural knowledge and interpretation, interactive skills, strategic adaptation, and critical cultural awareness. The Cronbach's α coefficients for the pre-test and post-test were 0.88 and 0.91, respectively. The oral task scoring scale has a total score of 100 points and includes five categories: informational completeness, pragmatic appropriateness, interaction progression, correction and negotiation, and cultural appropriateness. The immersion scale consists of 12 items with a Cronbach's α of 0.89; the cognitive load scale consists of 7 items with a Cronbach's α of 0.83.

To avoid masking dimensional differences with a single mean, this study further constructs a comprehensive ICC index for learners, as shown in Equation (2).

$$ICC_i = \sum_{k=1}^5 w_k z_{ik} \quad (2a)$$

$$\sum_{k=1}^5 w_k = 1 \quad (2b)$$

In the formula, ICC_i represents the comprehensive intercultural communication competence index of the i th learner; z_{ik} represents their standardized score on the k th dimension; and w_k is the dimension weight. This study applies equal weighting to the five dimensions, i.e., $w_k = 0.20$. To simultaneously measure "how well they learned" and "at what cost," this study further defines situational conversion efficiency E_i , as shown in Equation (3).

$$E_i = \frac{\lambda P_i + \mu R_i + \nu A_i}{1 + \eta L_i} \quad (3a)$$

$$\lambda + \mu + \nu = 1 \quad (3b)$$

In the equation, E_i represents the situational transfer efficiency of the i th learner; P_i denotes the oral task score; R_i denotes the reflection quality score; A_i denotes the cultural appropriateness score; L_i denotes the cognitive load level; and λ , μ , and ν represent the weights of the three types of outcome indicators, respectively. In this study, $\lambda = 0.40$, $\mu = 0.25$, and $\nu = 0.02$ were selected.

From the angle of statistics, this research firstly applied independent-sample t-tests and chi-square tests to verify the equality of the two baseline groups, then utilized mixed-design ANOVA to inspect the group \times time effect, and employed ANCOVA to make adjustment for post-test outcomes when the situation requires it. We carried out analysis on behavior logs with between-group comparison methods and correlation analysis methods, therefore interview data we utilized for explaining core results [21, 22]. For the purpose of finding out where the influences of contextual parts come from, four kinds of cut-down small tasks were added to the experiment group on the 10th to 11th weekends. These were put out by a Latin square rotating way, having four versions: "taking away cultural cues," "taking away branch dialogue," "taking away thinking prompts," and "taking away the existence of peers." These things afterwards were put into comparison with the complete version in every one of the groups. The connection among the experiment flow, group comparisons and the result interactive surface is displayed by Figure 3.

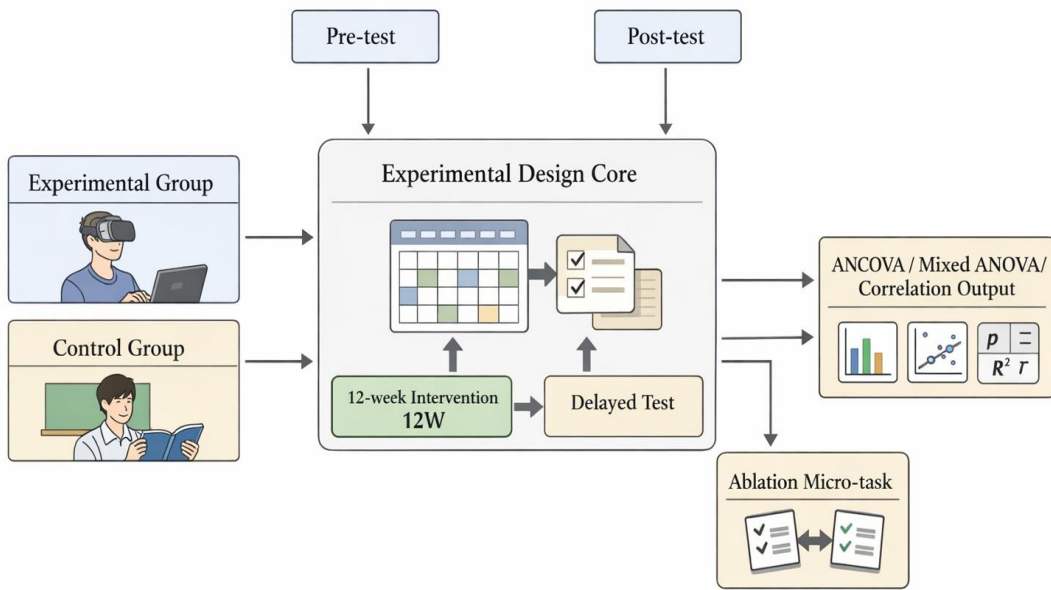


Figure 3: Schematic representation of the experimental protocol, group comparisons, and results interpretation interface.

In Figure 3, the experimental group and the control group have three evaluation time points—pre-test, post-test, and delayed test—this guarantees a consistent foundation for carrying out the comparison of differences between the two groups. The main difference is located in the middle intervention channel: the experiment group has a 12-week intervention which is followed by the adding of ablation small tasks in the later stage, therefore the contrast group only gets normal teaching arrangements. This protocol structure makes main effects, retention effects, and module effects can be analyzed step by step inside the same evaluation framework: the pre-test is utilized to consider initial differences, the post-test is for finding overall intervention effects, the delayed test is for checking skill retention, and the Ablation Micro-task is further used to divide the contributions of key components, for example, cultural cues, branching

dialogue, and feedback support. On the right part of the statistics result, data coming from the post-test, delayed test, and ablation micro-task are put together into ANCOVA, mixed ANOVA, and correlation analysis interfaces, which separately correspond to group main effects, time-interaction effects, structural differences, and mechanism explanations.

3 Results and Discussion

3.1 Overall Improvement Effects and Structural Differences

After completing sample control and implementing the intervention, the first question to address is: Did the VR scenario construction generally enhance students' cross-cultural communication competence, and in which competency dimensions was this enhancement primarily reflected? The overall results are shown in Table 2.

Table 2: Main effect outcomes on cross-cultural communication ability and spoken task completion

Metric	Experimental Group Pre-Test	Experimental Group Post-Test	Experimental Group Delay	Control Group Pre-Test	Control Group Post-Test	Control Group Delay	Group× Time F	p	ηp^2
ICC Total Score (1-5)	3.21±0.38	4.06±0.34	3.91±0.36	3.19±0.41	3.58±0.39	3.47±0.40	34.72	<0.001	0.219
Openness Attitude	3.34±0.42	4.18±0.37	4.06±0.38	3.31±0.40	3.71±0.41	3.62±0.43	21.85	<0.001	0.149
Cultural Knowledge and Interpretation	3.09±0.45	3.94±0.39	3.79±0.40	3.11±0.43	3.46±0.41	3.38±0.42	26.14	<0.001	0.173
Interaction Skills	3.17±0.41	4.11±0.35	3.97±0.36	3.16±0.39	3.55±0.40	3.43±0.41	31.09	<0.001	0.201
Strategy Adjustment	3.12±0.43	4.03±0.36	3.88±0.37	3.10±0.41	3.47±0.39	3.36±0.40	29.67	<0.001	0.192
Critical Cultural Awareness	3.32±0.40	4.02±0.38	3.87±0.39	3.28±0.39	3.70±0.40	3.56±0.41	18.94	<0.001	0.132
Speaking Task Score (0-100)	71.8±7.4	85.6±6.3	82.9±6.8	72.1±7.1	78.4±6.9	76.3±7.2	37.51	<0.001	0.238

Table 2 gives the changes of average scores for the experiment group and the control group in pre-test, post-test, delayed test, and also the group × time mutual influence effect. On the whole, the ICC total score of the experimental group was elevated from 3.21 to 4.06, and in the delayed test it still maintained at 3.91; the control group's score has a growth from 3.19 to 3.58 but after that it fell to 3.47. The interaction effect between Group and Time has attained statistical significance ($F = 34.72$, $p < 0.001$, $\eta p^2 = 0.219$), hence it indicates that the improvement and retention levels of the experimental group are significantly higher than the corresponding levels of the control group. The development tendencies of spoken task marks were in accordance with this: the experiment group's mark rose from 71.8 to 85.6, with a delayed repeated test mark of 82.9; In the control group, the scores increased from 72.1 points to 78.4 points, then they dropped to 76.3 points in the delayed test, therefore, the interaction effect is $F = 37.51$, $p < 0.001$, $\eta p^2 = 0.238$. This shows that the advantages of the VR scene intervention have gone past immediate classroom achievement to the short-term memory stage after the intervention ended.

Table 2 also reveals that the internal structure of the improvements was not uniform. The three dimensions with the largest increases in the experimental group were interactive skills, strategic adaptation, and open attitude. Specifically, interactive skills increased from 3.17 to 4.11, with a delayed test score of 3.97; strategic adaptation increased from 3.12 to 4.03, with a delayed test score of 3.88; and open-mindedness rose from 3.34 to 4.18, with the delayed test

score remaining at 4.06. In contrast, while cultural knowledge and interpretation, as well as critical cultural awareness, also showed significant improvements, their increases were slightly lower. This suggests that the direct advantages of the VR scenario are primarily reflected in dynamic abilities such as interactive advancement, responsive correction, and relational adjustment, whereas dimensions related to knowledge comprehension tend to show more steady improvements. The control group also has moderate growth on all five dimensions, but the scores of post-test and delayed test are generally lower than those of the experimental group. The difference was especially obvious in interaction abilities and strategy adjustment, hence it indicates that traditional situation-based teaching can promote basic understanding but it has difficulty in helping high-density, multi-turn communication interactions. To further examine the temporal trajectories corresponding to Table 2, Figure 4 displays the changes in total ICC scores and individual dimensions across the three time points.

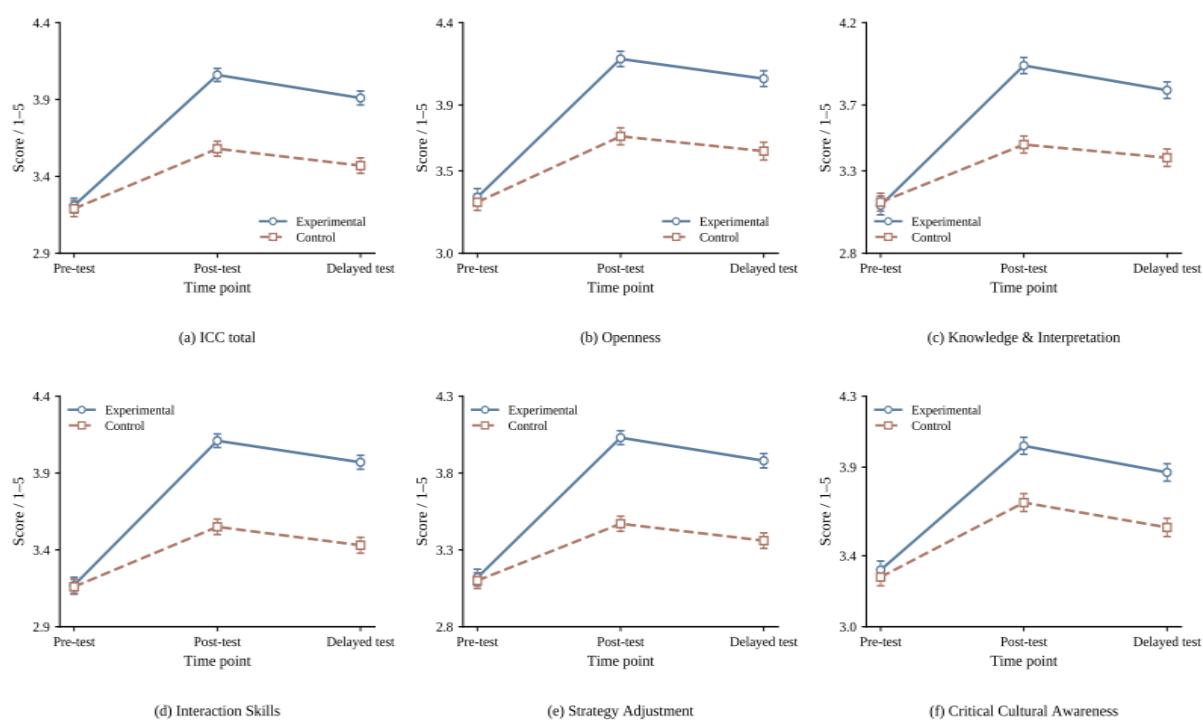


Figure 4: Pre-test-Post-test-Delayed-test change curves for the ICC dimensions between the experimental and control groups.

In Figure 4, the experiment group showed a consistent tendency on both the ICC total score and the five dimensions: a significant rise from pre-test to post-test, then a small decrease at the delayed test but still staying higher than baseline levels. By way of comparison, the curve of the control group possesses a more gradual incline, hence the descent in the delayed period is more obvious. The curve of ICC total score can most well explain this difference: the group of experiment reached a higher flat stage in the post-test period; although the delayed test had a decline, it still was significantly higher than the post-test level of the control group. Among the five dimensions, the curves of interactive skills and strategic adaptation have the biggest difference between groups, hence this shows that the multi-round interactions, role-related feedback, and cultural clues in the VR environment have a more direct influence upon these two abilities. The promotion of openness was comparatively stable, hence it indicates that the immersed environment has a continuous function in lowering communication defensiveness and increasing willingness to take part.

The combination of Table 2 and Figure 4 makes us discover that the influence of VR scene

building on immersive English teaching is not averagely distributed on all ability dimensions, instead, it puts the enhancement of interaction abilities which connect most tightly with real-world communication in the first place. In this circumstance, students not only made more speaking contents but also promoted their correction and communication abilities via more frequent answers to conversation partners and cultural evaluations. This result offers a basis for the situation-related contrasts which are talked over below.

3.2 Scenario-Specific Performance, Interactive Behavior, and Case Analysis

Having established the overall improvement, the next step is to further distinguish how these advantages are distributed across different scenarios and to determine whether behavioral-level changes can explain the differences between groups. The task scores, response latency, and correction success rates of the two groups across the five scenarios are shown in Table 3.

Table 3: Task Completion Degree, Reaction Time-Lag, and Correcting Success Proportions Under Various Situations

ontext	Experimental Group Task Score	Control Group Task Score	Experimental Group Response Time (s)	Control Group Response Time (s)	Experimental Group Correction Success Rate (%)	Control Group Correction Success Rate (%)	<i>p</i>
International Student Reception	84.2±6.1	78.1±6.8	1.72±0.31	2.08±0.36	75.4	63.1	0.002
Homestay Dinner	86.9±5.7	79.4±6.5	1.58±0.28	2.17±0.40	79.6	64.8	<0.001
Academic Seminar Q&A	83.5±6.6	77.0±7.2	1.81±0.34	2.24±0.42	72.3	60.9	0.004
Public Service Processing	85.7±6.0	78.6±6.7	1.64±0.29	2.11±0.37	77.8	62.4	0.001
Intercultural Project Negotiation	88.1±5.4	79.8±6.2	1.49±0.26	2.05±0.35	81.5	62.2	<0.001

With respect to task scores in Table 3, the experiment group has better scores than the control group on all five situation cases. The maximum score has been obtained in cross-cultural project negotiation, which is 88.1 points; that is followed by the dinner held by the host family, at 86.9 points; public service procedures and international student orientation scored 85.7 and 84.2 points, respectively; and the academic seminar Q&A scored relatively lower at 83.5 points. The control group's scores for the corresponding scenarios were 79.8, 79.4, 78.6, 78.1, and 77.0, respectively, which were lower overall than those of the experimental group, with the gap being more pronounced in the project negotiation and homestay scenarios. In terms of response latency, the mean for the experimental group across the five scenarios was 1.65 s, compared to 2.13 s for the control group, indicating that the experimental group was able to generate effective responses more quickly after entering the interaction. Regarding correction success rates, the experimental group's values ranged from 72.3% to 81.5% across the five scenarios, while the control group's ranged from 60.9% to 64.8%, indicating that the experimental group not only responded faster but also more readily corrected their expressions during the interaction. Table 3 shows that the benefits of the VR scenarios are not evenly distributed but are more concentrated in tasks requiring higher levels of relationship negotiation and cultural judgment. To further examine the coupling relationship between task performance and response latency across different scenarios, the changes in oral performance and response latency across the five scenarios are shown in Figure 5.

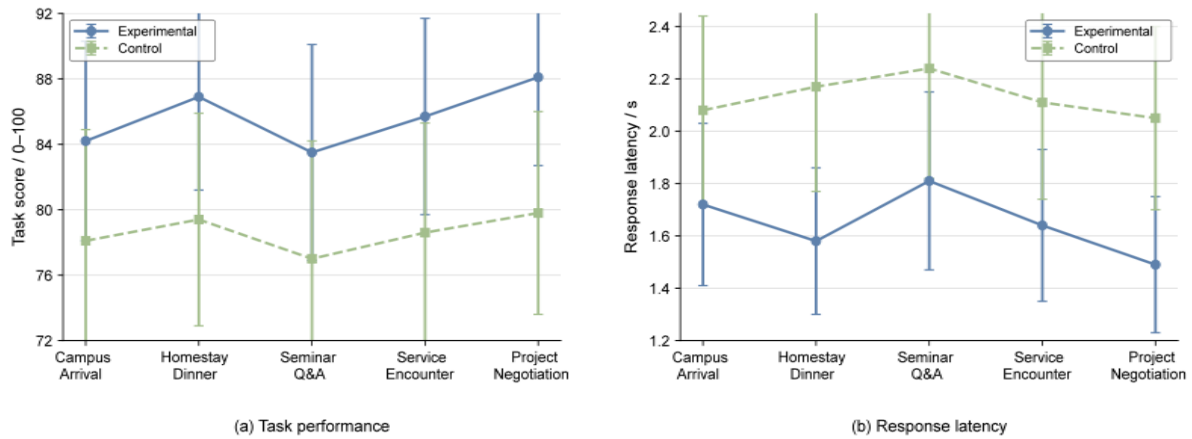


Figure 5: Coupling curves of oral performance and response latency across five VR scenarios.

In Figure 5, the experiment group continuously showed a "higher point-faster reaction time" pattern in all five situations, while the comparison group on the whole presented "lower points-slower reaction times." Notably, the greatest intergroup divergence was observed in the curves for cross-cultural project negotiation and the host family dinner scenarios, indicating that these two scenarios best demonstrate the advantages of immersive environments. In the project negotiation scenario, the experimental group not only achieved the highest scores among all groups but also reduced response latency to 1.49 s, indicating that learners were able to organize their positions and select strategies more quickly when dealing with the expression of disagreements, budget coordination, and compromise negotiations. In the homestay dinner scenario, the experimental group also demonstrated higher scores and shorter response times, indicating that everyday communication tasks involving politeness regulation, relationship maintenance, and implicit cultural norms can similarly benefit from direct support provided by the VR environment. In contrast, although the academic seminar Q&A task still outperformed the control group, both the curve slope and the intergroup gap were relatively gentle, suggesting that this scenario is more heavily influenced by conceptual retrieval load and immediate organizational pressure, resulting in a slower release of benefits compared to negotiation- and politeness-oriented tasks. To explain the scenario differences shown in Figure 5, the standardized intensity distribution of key interactive behaviors is presented in Figure 6.

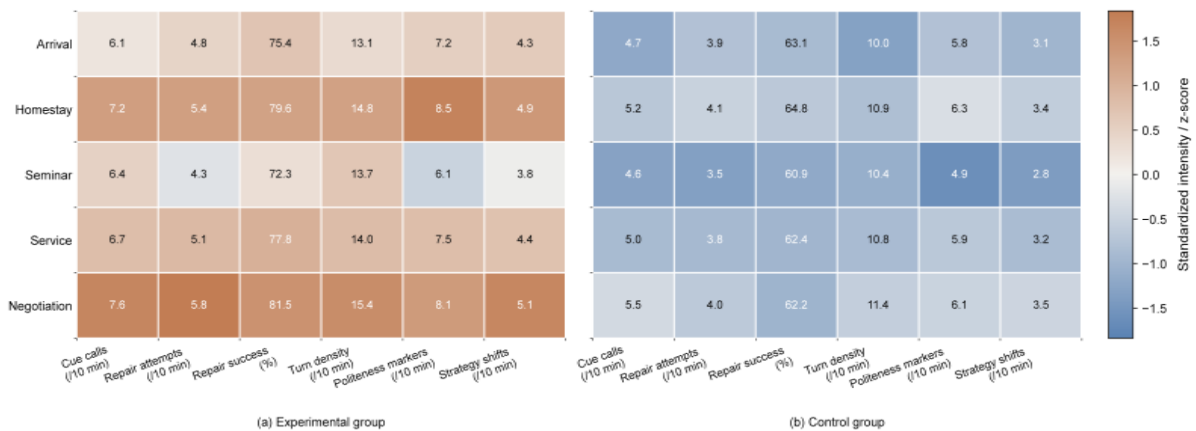


Figure 6: Heatmaps of cultural cue invocation, correction strategies, and interaction density.

In Figure 6, the experiment group did better than the control group on all behavior measurement indexes, these include cultural cue calling, correction trying, successful correction,

interaction turn density, politeness mark usage, and strategy changing strength. After being calculated as standardized frequency within every 10 minutes, the mean intensity of cultural cue calling in the experimental group was 6.8, as compared with 5.0 in the control group; the average number of interaction turns was 14.2 in the experimental group, and 10.7 in the control group; and the average successful rate of correction was 77.3 percent in the experimental group and 62.7 percent in the control group. Further deeper analysis makes it clear that the correction successful rate and interaction density are the highest in the project negotiation situation, achieving 81.5% and 15.4 rounds per 10 minutes, each one, therefore showing this situation most effectively brought about multi-round negotiation and continuous correction. The dinner scene of host family displayed the most obvious usage of politeness labels and cultural hints, achieving 8.5 and 7.2 times every 10 minutes, separately, hence this shows that daily relationship interactions depend more on the acknowledgment of cultural rules and the usage of expressive buffering tools.

Representative examples further let us know the special nature of this change. Taking the project negotiation work of experimental group member E17 as an example, in the pre-test stage, this student usually directly refused the opposite side's proposal when confronting opposite opinions, lacked buffer words and replacement proposals in his words, and the negotiation process often ended in two rounds. In the stage after the test, their speaking order was adjusted to first express recognition of the opposite side's viewpoint, then give an account of budget limits, next put forward a mutual concession, and at last make sure of the team's common opinion through questions that are asked afterwards. This change has not only reflected an increase of language size, but also a progress in strategic choosing, relation dealing, and evaluation of cultural appropriateness. By comparison, the alike students in the control group, even though they can give more information points on the post-test, still have the tendency to reach personal conclusions fast once a disagreement appears, which shows that their negotiation flexibility and revision depth are weaker.

3.3 Module Ablation, Efficiency Analysis, Sources of Error, and Implications for Deployment

After confirming the effectiveness of the scenario, it is necessary to further distinguish: which specific components of the scenario account for the primary gains, and whether these benefits justify the deployment costs. To this end, this study implemented four types of stripped-down microtasks for the experimental group and compared the results with those of the full version. The results of the key module ablation are shown in Figure 7.

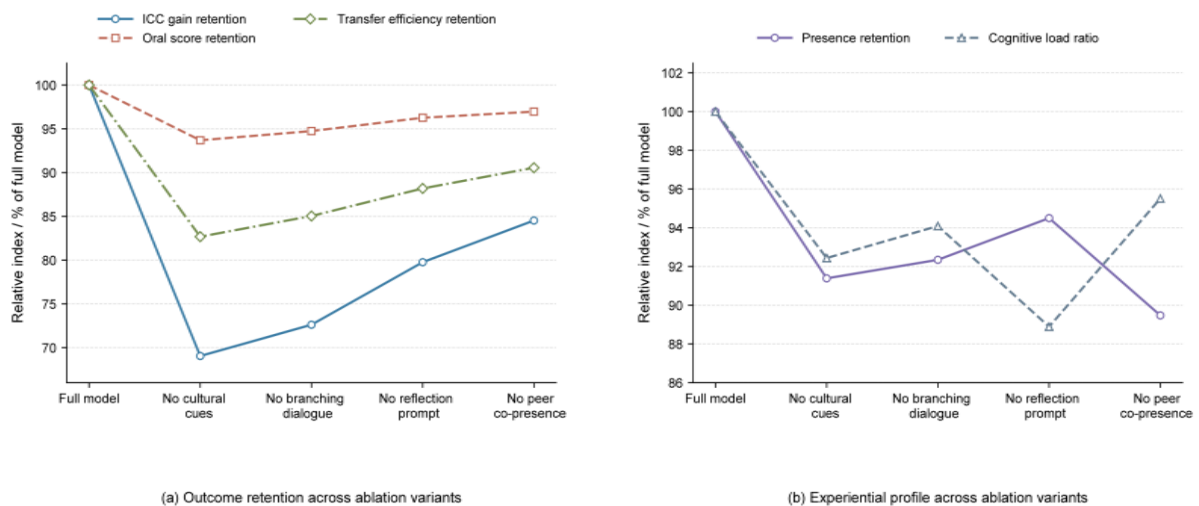


Figure 7: Curves showing the impact of ablating key modules in " " on learning results and load metrics.

In Figure 7, the ablation results indicate that the cultural cue layer and branch dialogue are the two most critical modules. The comprehensive task score for the full version is 85.6; it drops to 80.2 after removing the cultural cue layer, to 81.1 after removing branch dialogue, to 82.4 after removing reflective prompts, and to 83.0 after removing peer presence. Calculated according to Equation (2), the gain index of ICC_i for the full version is 0.84; after removing the cultural cue layer and branch dialogue, it drops to 0.58 and 0.61, respectively, showing the most significant decline. Calculated using Equation (3), the situational transformation efficiency E_i of the full version is 1.27, while the four abridged versions drop to 1.05, 1.08, 1.12, and 1.15, respectively. This indicates that once students enter the VR scene, what truly drives cross-cultural learning outcomes is not merely "seeing the scene," but rather "having to process cultural cues within the scene and make choices among different response paths." The relationship between presence, cognitive load, and learning gains is shown in Figure 8.

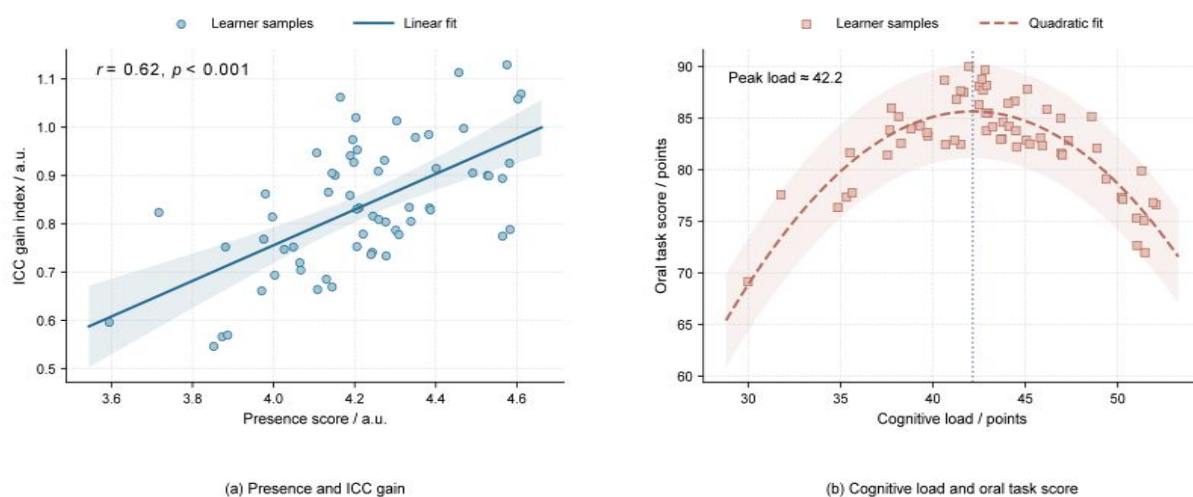


Figure 8: Scatter plot of the relationship between presence, cognitive load, and learning gains.

The relationship analysis in Figure 8 further explains the benefit threshold. The mean immersion score for the experimental group was 4.18/5, showing a significant positive correlation with the ICC_i gain ($r = 0.62$; $p < 0.001$); Cognitive load exhibited an inverted U-shaped relationship with oral task scores, with the optimal range falling between approximately 38 and 45 points; when the load exceeded 52 points, both expressive accuracy and cultural appropriateness declined simultaneously. This implies that "high immersion" in VR instruction does not automatically equate to "high learning gains"; moderate task pressure and a clear feedback structure are more critical. Recent studies on the balance between interaction intensity and learning load have reached similar conclusions [23].

From the perspective of deployment efficiency, for the experimental group, the average device preparation time for each student dropped from 4.6 minutes in the first week to 2.1 minutes in the fourth week, and it was stabilized at 1.8 minutes from that time onwards; the proportion of reports on headgear uncomfortable feeling was 7.9%, mainly concentrated on the opening two weeks. Errors mainly came from four respects: firstly, some students still depended on written-down words in high-pressure environments, hence leading to not enough deepness of real negotiation; Second, the high cognitive burden which is linked to concept taking-out in academic talk situations has limited the space for giving cultural judgments; Third, although the existence of peers made interaction density higher, it therefore also brought extra anxious

feeling to a number of students; fourth, hence if instructors do not in time check learning records and give pointed direction, the advantages of the reflection and feedback layer hence were greatly reduced.

Therefore, the deployment of VR in immersive English instruction should not merely pursue "device coverage"; rather, it requires the establishment of stable course units centered around scenario libraries, cultural cue databases, and feedback templates. Recent XR reviews and meta-analyses have consistently pointed out that the educational value of XR is highly correlated with the quality of task design, teachers' instructional design capabilities, and technical proficiency [24, 25]. Applied to the findings of this study, this conclusion can be specified as follows: only when cultural cues, branching interactions, and reflective feedback are integrated into a single instructional unit can VR transition from a demonstrative medium to an argumentative and training-oriented teaching tool.

4 Conclusion

For the problems of not enough situational reality, not strong cultural hints, and not steady cross-cultural shift in immersion English teaching, this research built a VR-based situational teaching frame for university classrooms, and verified its teaching effects through quasi-experiments, delayed tests, and multi-source data analysis. The outcomes show that this framework is able to condense language tasks, cultural judgments and interactive corrections into one single situational unit, hence therefore enhancing the true communication direction of English study and the manifestation of cross-cultural abilities.

(1) This research has built a framework for the arrangement of situations and the assessment of study results in English immersion teaching. Taking five high-frequency tasks as the core-campus direction guide, homestay living experiences, academic question answering, public service items, and project business negotiations-the study has built VR scenario units which include cultural hints, branched dialogues, same-time existence of peers, teacher framework support, and reflective feedback information. Through the combining of scale tables, spoken assignments, behavior records, and interview materials into one united assessment frame, this research has offered a steady base for the explanation of multi-source proofs inside the classroom.

(2) Experimental results confirm that this framework has a sustained positive effect on intercultural communication competence and oral task performance. The experimental group's ICC total score increased from 3.21 to 4.06, with the delayed retest score remaining at 3.91, while the control group's post-test and delayed retest scores were 3.58 and 3.47, respectively; the experimental group's oral task scores increased from 71.8 to 85.6, with a delayed test score of 82.9, whereas the control group's corresponding scores were only 78.4 and 76.3. Scenario-specific results showed that task scores for cross-cultural project negotiation and host family dinner reached 88.1 and 86.9, respectively. The experimental group's average response latency decreased to 1.65 s, and the correction success rate increased to 77.3%, indicating that the advantages of the VR scenario are primarily evident in tasks requiring multiple rounds of interaction, relationship management, and cultural judgment. Ablation results further indicate that the cultural cue layer and branch dialogues are key sources of performance gains; the scenario conversion efficiency of the full version was 1.27, which decreased to 1.05 and 1.08, respectively, after removing the cultural cue layer and branch dialogues.

(3) This study still has limitations, including a single source of samples, a limited intervention period, and device conditions constrained by classroom resources. Future research could be expanded to include samples from multiple schools, long-term tracking, and adaptive dialogue agent environments, and continue to examine transfer stability and deployment cost

control across different English courses in various disciplines.

References

- [1] Byram, M. (2021). Teaching and assessing intercultural communicative competence revisited. Bristol: Multilingual Matters.
- [2] Deardorff, D. K. (2006). Identification and assessment of intercultural competence as a student outcome of internationalization. *Journal of Studies in International Education*, 10(3), 241-266.
- [3] Shadieff, R., & Sintawati, W. (2020). A review of research on intercultural learning supported by technology. *Educational Research Review*, 31, 100338.
- [4] Dhimolea, T. K., Kaplan-Rakowski, R., & Lin, L. (2022). A systematic review of research on high-immersion virtual reality for language learning. *TechTrends*, 66, 810-824.
- [5] Luo, S., Zou, D., & Kohnke, L. (2024). A systematic review of research on xReality (XR) in the English classroom: Trends, research areas, benefits, and challenges. *Computers & Education: X Reality*, 4, 100049.
- [6] Yudintseva, A. (2023). Virtual reality affordances for oral communication in English as a second language classroom: A literature review. *Computers & Education: X Reality*, 2, 100018.
- [7] Rahmanu, I. W. E. D., & Molnár, G. (2024). Multimodal immersion in English language learning in higher education: A systematic review. *Heliyon*, 10(19), e38357.
- [8] Yan, W., & Lowell, V. L. (2025). The evolution of virtual reality in foreign language education: From text-based MUDs to AI-enhanced immersive environments. *TechTrends*, 69(4), 853-858.
- [9] Li, Q., Liu, Q., Jiang, M., et al. (2024). Virtual reality as a bridge: Enhancing intercultural learning and communication skills among international students. *International Communication of Chinese Culture*, 11, 385-411.
- [10] Knutzen, K., Rothenberger, L., Tribusean, I., et al. (2025). Using social virtual reality in teaching intercultural communication. *Technology, Knowledge and Learning*, 30(2), 1167-1187.
- [11] Tafazoli, D. (2024). From virtual reality to cultural reality: Integration of virtual reality into teaching culture in foreign language education. *Journal for Multicultural Education*, 18(1-2), 6-24.
- [12] Hsu, L. (2024). Exploring EFL learners' acceptance and cognitive absorption at VR-based language learning: A survey and experimental study. *Heliyon*, 10(3), e24863.
- [13] Ding, M. (2024). The impact of high-immersion virtual reality on EFL learners' foreign language speaking anxiety: A mixed-method approach. *ReCALL*, 36(3), 287-305.
- [14] Yudintseva, A. (2024). An exploration of low- and high-immersive virtual reality

- modalities for willingness to communicate in English as a second language. *Computers & Education: X Reality*, 5, 100076.
- [15] Gu, L. (2025). Beyond the classroom - exploring the impact of virtual reality exposure on foreign language anxiety with the mediating role of ESL Chinese learners' communicative confidence and fluency. *Humanities and Social Sciences Communications*, 12, 1349.
- [16] Makransky, G., & Petersen, G. B. L. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*, 33, 937-958.
- [17] Miguel-Alonso, A., Viera, A., Lima, J., et al. (2023). Countering the novelty effect: A tutorial for immersive virtual reality learning environments. *Applied Sciences*, 13(1), 593.
- [18] Miller De Rutté, A. M. (2024). Virtual reality, medical Spanish, and the L2 motivational self system: A case study. *Technology, Knowledge and Learning*, 29(2), 827-841.
- [19] Miller De Rutté, A. M. (2024). The design and implementation of virtual reality simulations in Spanish for the health professions: What are learners' perceptions? *Computer Assisted Language Learning*, 1-21.
- [20] Luan, L., Hwang, G. J., Yi, B., et al. (2025). The effects of a self-developed virtual reality environment on college EFL learners' vocabulary learning. *Interactive Learning Environments*, 33(1), 335-346.
- [21] Shadiey, R., Wang, X., & Shen, S. (2025). Effects of immersion and interactive strategies on students' intercultural competence in virtual learning environments. *Education and Information Technologies*, 30(5), 5883-5919.
- [22] Wang, X., Shadiey, R., Sintawati, W., et al. (2025). A comparative study on the impact of four different SVVR-supported intercultural learning environments on learners' intercultural competence and learning engagement. *Interactive Learning Environments*, 33(4), 2887-2913.
- [23] Cai, L., Wang, Z., Wang, C., et al. (2025). Interactivity and signaling in immersive virtual reality: Effects on EFL learning experiences and outcomes. *Computers & Education*, 238, 105412.
- [24] Christou, E., Parmaxi, A., & Christoforou, M. (2025). Implementation and application of extended reality in foreign language education for specific purposes: A systematic literature review. *Universal Access in the Information Society*, 24, 2061-2076.
- [25] Kumar, A., & Gorai, J. (2025). Effectiveness of augmented reality and virtual reality interventions on learning outcomes: A meta-analysis in higher education. *TechTrends*, 69, 1207-1220.