Effectiveness of Immersive Contextual English Teaching Based on Fuzzy Evaluation

Mei Niu*

Basic Courses Department, Jiyuan Vocational and Technical College, Jiyuan 459000, Henan, China

***Abstract*—Investigating the real-world impact of immersive contextual instruction on English language education, verifying its contribution to the enhancement of linguistic skills and the improvement of learning attitudes, and evaluating the practicality and worth of fuzzy evaluation in gauging teaching efficacy. A fuzzy complete assessment model was built utilizing the language competency test and the learning attitude questionnaire, and the teaching effect was quantitatively examined based on the experimental data using methods such as affiliation function and weight calculation. The study's findings revealed that students in the experimental group performed much better than students in the control group in terms of language competence and learning attitudes, with an overall fuzzy score of 88.5 compared to 74.8 in the latter. The statistical test indicated a significant difference between the groups (p<0.001). The study also confirmed the scientific and practical validity of fuzzy evaluation in the assessment of multidimensional educational efficacy. Immersion contextual English teaching provides considerable benefits for improving students' language skills and learning attitudes. The fuzzy assessment method introduces a new instrument for quantitative research on teaching efficacy and has a wide range of potential applications.**

***Keywords*—Fuzzy evaluation; immersion; contextual English teaching; teaching effectiveness; teaching assessment**

I. Introduction

With the acceleration of globalization, the improvement of the teaching effect of English, as an important tool for international communication, has increasingly become the focus of research in the field of education [1]. The traditional English teaching mode is mostly based on teachers explaining and students listening to lectures, which lacks contextual immersion, resulting in students' language application ability often being difficult to effectively improve [2]. In recent years, immersion teaching, as an innovative teaching mode, has gradually gained the attention of more and more educators [3]. By simulating the real language use environment, immersion teaching enables students to experience language learning in real situations, thus improving their language proficiency and cultural understanding [4]. However, although the immersion mode of teaching shows obvious advantages in English language teaching, the evaluation of its teaching effectiveness faces a big challenge [5]. Traditional assessment methods often rely on quantitative test scores or qualitative teacher assessment, but these methods are difficult to fully reflect students' comprehensive language proficiency in immersion contexts [6]. Therefore, how to scientifically and accurately evaluate the effects of immersion context teaching has become an important topic in current educational research.

In the complex and dynamic educational context, the fuzzy evaluation method, as a tool adept at handling uncertainty and vagueness, demonstrates its distinctive merits [7]. It can integrate both qualitative and quantitative metrics and conduct a holistic analysis of teaching outcomes from various evaluation perspectives, thereby offering a novel approach for assessing the efficacy of immersive English instruction [8]. Against this backdrop, this study aims to employ the fuzzy evaluation method to carry out a systematic investigation into the teaching effectiveness of immersive contextual English, with the goal of furnishing educators with a more precise and all-encompassing teaching effectiveness assessment instrument [9]. The primary objective of this research is to construct a fuzzy evaluation model suitable for immersive contextual English teaching via empirical research, identify the crucial factors influencing teaching effectiveness, and analyze the data through specific cases [10]. It is anticipated that this research will provide theoretical backing for future English teaching endeavors and an effective basis for decision-making for educational policymakers.

The content structure of this paper is divided into five main sections. Section II of the research review comprehensively combed the domestic and international literature related to this study, including the theoretical foundation of immersive contextual English teaching and the progress of applied research, while analyzing the characteristics of the fuzzy evaluation method and its application value in educational research. Then, Section III of the research method elaborates the design framework of this study in detail, including the selection of experimental samples, data collection methods, the construction process of the fuzzy comprehensive evaluation model, and the specific calculation steps. Subsequently, Section IV of the results and discussion presents the differences in teaching effects between the experimental group and the control group based on empirical data and combines the quantitative results of the fuzzy evaluation with an in-depth discussion of the superiority of immersive contextual teaching and the specific performance of its teaching effects. Finally, Section V of the conclusion summarizes the main findings of the study, analyzes the shortcomings of the study, and looks forward to the future research direction. The parts are interlocked and work together to serve the achievement of the research objectives.

II. Research Review

To fully understand the theoretical background and the current application status of immersive contextual teaching and

*Corresponding Author

fuzzy evaluation, this part will be developed from the following two aspects: the current research status of immersive contextual teaching and its evaluation methods, and the progress of the application of fuzzy evaluation methods in the field of education and its advantages.

A. Research Status and Evaluation Methods of Immersive Contextual Teaching and Learning

The core of the immersive contextual teaching model is to enhance learners' language communication skills by constructing real or simulated language use contexts so that they can practice the language in a near real environment [10]. This model has been widely used in language learning, science experiments cultural courses, etc. Its advantage is that it can enhance students' participation and language acquisition. Some studies have shown that students' language expression and cultural understanding are significantly improved when immersive learning environments are constructed through virtual reality (VR) technology or classroom scenario simulation [11]. However, existing studies have mostly focused on the following three ways of evaluating the effects of immersive teaching: (1) performance assessment based on quantitative data, such as standardized test scores, and comparison of teaching effects through pre-tests and post-tests. (2) Qualitative assessment, such as evaluating teaching effectiveness through students' self-feedback, interviews, or teachers' classroom observations. (3) Mixed assessment, i.e., combining quantitative and qualitative indicators, such as combining test scores and questionnaires for comprehensive analysis. Although these methods can reflect teaching effectiveness to a certain extent, they are often inadequate in dealing with complex and multidimensional evaluation issues [12]. For example, it is difficult to measure the comprehensiveness of students' language useability by relying only on test scores, while subjective interviews and observations are susceptible to the subjective bias of the evaluator [13]. Table I presents for typical immersion teaching evaluation methods and their strengths and weaknesses; therefore, there is a need to introduce evaluation methods that are more scientific, multidimensional, and capable of quantifying complex phenomena.

TABLE I TYPICAL IMMERSION EVALUATION METHODS AND THEIR ADVANTAGES AND DISADVANTAGES

Evaluation Methodology	Vintage	Drawbacks
quantitative assessment	(1) Strong objectivity and easy access to and analysis of data (2) Can be used for comparison and statistical analysis of large samples	(1) Confined to numerical results, it is difficult to fully reflect the effectiveness of teaching and learning (2) Inability to capture learners' subjective experiences and emotional changes
Qualitative assessment	(1) The ability to dig deeper into learners' subjective feelings about teaching and learning (2) Helps capture details and dynamic changes in language use	(1) Highly subjective, with results susceptible to the personal biases of the evaluators (2) A small sample size makes it difficult to generalize to a wider area
Blended assessment	(1) Combines the strengths of quantitative and qualitative methods (2) Teaching effectiveness can be evaluated more comprehensively from multiple dimensions	(1) Complexity of data processing, which may increase the cost of the study (2) The qualitative component remains vulnerable to subjectivity
Assessment based on fuzzy evaluation	(1) Suitable for dealing with complex, multidimensional, and difficult-to-quantify problems (2) Ability to synthesize different indicators and generate overall evaluations (3) Reducing the bias of single data	(1) The model-building process may be subjective in terms of parameter setting (2) The affiliation function and weights should be set scientifically and carefully.

B. Application of Fuzzy Evaluation Methods in Education

The fuzzy evaluation method originated from the fuzzy set theory proposed by Zadeh, which aims to solve the limitations of traditional evaluation methods in dealing with vagueness and uncertainty. In the field of educational evaluation, fuzzy evaluation has attracted much attention because of its ability to integrate multidimensional indicators and handle quantitative and qualitative data [14]. In recent years, fuzzy evaluation methods have been widely used in the fields of teaching quality evaluation, students' comprehensive quality evaluation, and course satisfaction analysis. A study applied the fuzzy comprehensive evaluation method to the assessment of university teaching quality and established a set of multi-dimensional comprehensive evaluation models by setting subjective weights and affiliation functions, which greatly improved the scientificity and persuasiveness of the evaluation results. Similarly, another study proved the high applicability of the fuzzy evaluation model in the assessment of multi-indicator teaching effectiveness in the study of vocational skills teaching. Compared with traditional evaluation methods, fuzzy

evaluation has several advantages [15]. The first is the integration of multidimensional data. It can synthesize and analyze data of multiple dimensions. Then it deals with fuzziness and subjectivity. Through the affiliation function, qualitative evaluations are quantified into manageable mathematical models [16].The fuzzy evaluation method also exhibits significant adaptability in handling intricate and variable teaching scenarios. When integrated with the features of immersive contextual teaching, assessing its outcomes typically entails a multitude of complex and multidimensional factors (such as language proficiency, cultural comprehension, emotional attitudes, etc.), aligning well with the fundamental attributes of fuzzy evaluation [17]. Hence, utilizing fuzzy evaluation for gauging the effectiveness of immersive contextual English instruction can address the limitations of conventional assessment techniques and offer innovative perspectives for teaching effectiveness research. Fig. 1 delineates the construction phases of the fuzzy evaluation model, depicting the entire procedure from identifying evaluation indicators to executing the fuzzy synthesis operation. This aids users in comprehending how to apply the fuzzy evaluation approach to assess the impact of immersive contextual teaching [18]. Initially, it is crucial to establish the evaluation indicator system and choose the evaluation dimension that resonates with the attributes of immersive contextual teaching [19]. Subsequently, the membership function for each evaluation indicator is ascertained to convert qualitative assessments into a quantifiable range of values [20]. Thereafter, the weight of each evaluation indicator is determined based on expert opinions, student feedback, or pertinent literature [21]. Building on this, the fuzzy comprehensive evaluation matrix is formulated by integrating the membership function and weight of each indicator, and the membership values of each evaluation object across each dimension are aggregated into a matrix. Following this, the constructed fuzzy comprehensive evaluation matrix is manipulated to derive the overall score for each evaluation object. Lastly, in accordance with the operational outcomes, the final evaluation results are generated and can be displayed in formats such as scores or grades, assisting decision-makers in making informed judgments.

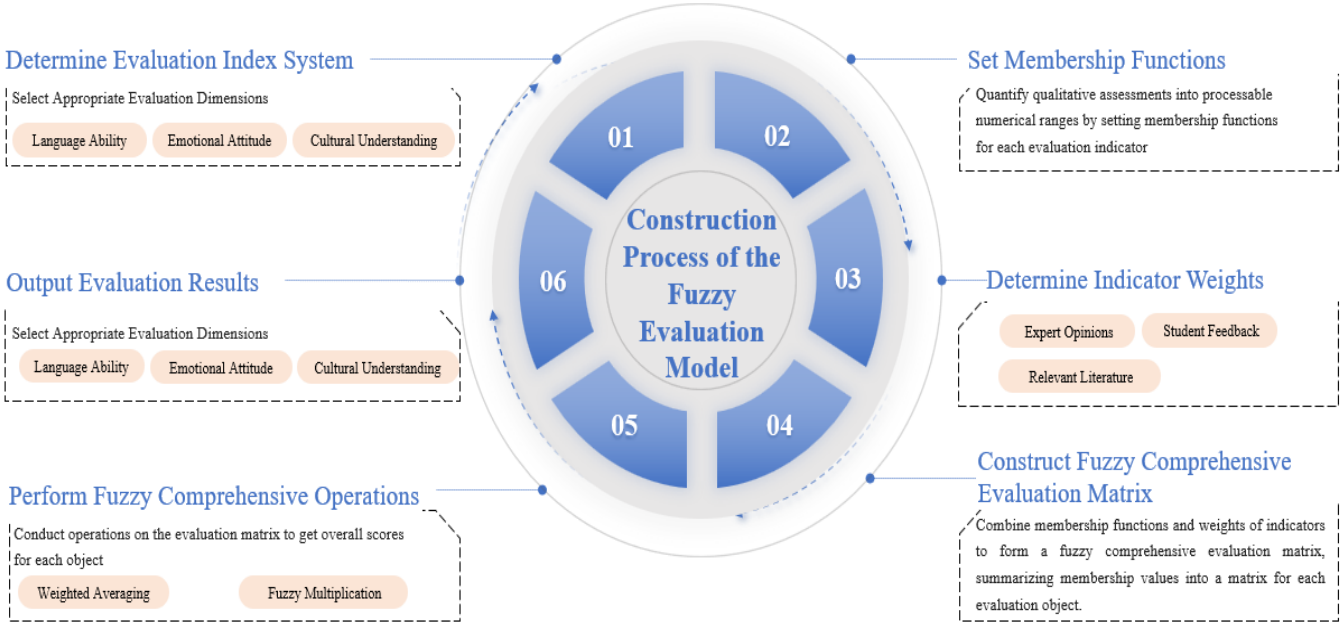

Fig. 1. Construction process of fuzzy evaluation model.

C. Review of the Study and Innovations

To summarize, immersive contextual teaching has received widespread attention in recent years due to its remarkable language-learning effect, but there are still major limitations in the evaluation methods of teaching effectiveness; fuzzy evaluation methods have opened up new paths for educational evaluation research due to their ability to handle multidimensional and complex data and their flexibility [22]. However, few studies have organically combined the two and conducted a comprehensive empirical analysis of actual cases in immersive contextual teaching. A comparison of the innovations of this study with previous studies is shown in Table II. Based on the above analysis, the innovations of this paper are mainly reflected in the following two aspects. (1) Introducing the fuzzy evaluation method into the assessment of the effect of immersive contextual teaching and constructing a multidimensional and highly adaptable fuzzy evaluation model. (2) Through empirical research and data analysis, the validity of the model is verified from actual cases to fill the shortcomings of existing research.

TABLE II COMPARISON OF THE STUDY'S INNOVATIONS WITH PREVIOUS STUDIES

Research	Previous Study	Innovative Points of This Study
Evaluation methods for immersive contextualized instruction	(1) Most studies use traditional quantitative assessments (e.g., achievement assessments, comparison of standardized test scores) (2) or qualitative assessment (e.g., interviews, classroom observations)	Introducing the fuzzy evaluation method into the assessment of immersive contextual teaching, breaking through the limitations of the traditional single method, and constructing a multi-dimensional and highly adaptable fuzzy evaluation model
Comprehensive assessment of teaching effectiveness	Most of the existing research focuses on single dimensions or quantitative indicators, such as academic performance or student feedback, and lacks an integrated approach to assessment.	Based on fuzzy evaluation, a comprehensive assessment of teaching effectiveness is realized through a comprehensive analysis of multiple evaluation dimensions (e.g. language proficiency, cultural understanding, affective attitudes, etc.).
Empirical research and data analysis	Few studies have empirically analyzed real-world cases for immersive contextual instruction.	This study combines real cases with in-depth empirical research to verify the effectiveness of the fuzzy evaluation model in immersive contextual teaching through data analysis.
Evaluation of Teaching Effectiveness Flexibility and Adaptability of	Traditional evaluation methods are often difficult to deal with in complex and dynamically changing teaching and learning contexts.	Fuzzy evaluation methods are flexible and adaptable to complex, dynamic teaching and learning contexts, and are highly adaptable, especially for assessing the effectiveness of immersive contexts.

III. METHODOLOGY

To explore the application of the fuzzy evaluation method in the assessment of English teaching effectiveness in immersive contexts, this study adopts an empirical-based research methodology to construct and validate the fitness model through case studies and data analysis. This section will specify the research design, sample selection, data collection and processing, construction of the fuzzy evaluation model, and its calculation process.

A. Study Design

This study adopts a mixed research method, combining qualitative and quantitative analysis to ensure both an in-depth analysis of the complex teaching phenomenon and the scientific and persuasive nature of the data results. The flow chart of the research design is shown in Fig. 2. The whole research process is divided into four main stages. The first is the contextual teaching implementation phase, in which a school's English course is selected and virtual reality technology is introduced to create multiple simulated learning scenarios, such as ordering food, medical conversations, and international conferences, to provide students with an immersive learning experience. Then comes the data collection phase, in which multidimensional data including standardized language test scores, classroom observation records, student interviews, and questionnaire feedback are collected after the implementation of teaching [23]. The third stage is the fuzzy evaluation model construction, which establishes a suitable fuzzy comprehensive evaluation model to quantify and comprehensively analyze the collected multidimensional data according to the research questions and data characteristics [24]. The final stage is the data analysis and validation stage, which uses tools such as Matlab and Python to process the data, calculate the specific scores of the fuzzy evaluation model, and carry out model adaptation validation [25]. This flowchart provides clear steps and directions for the research and helps to systematically conduct and analyze the research.

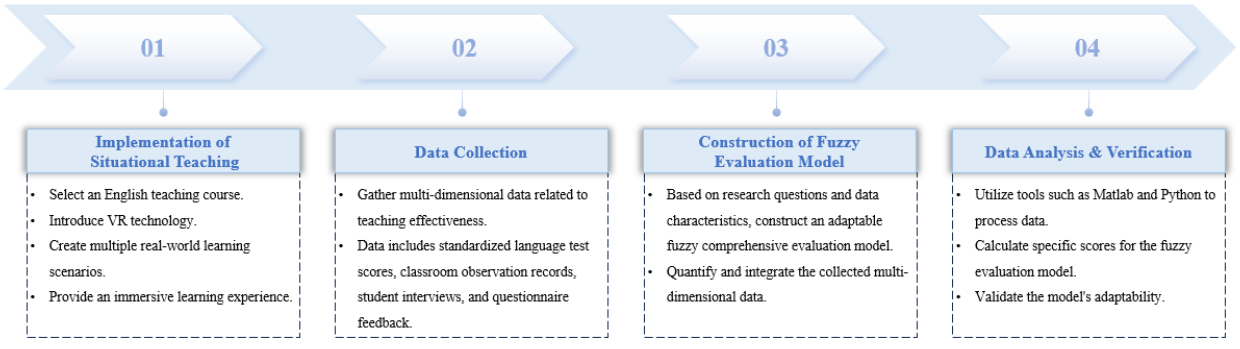

Fig. 2. Flowchart of the study design.

B. Sample Selection and Background

The participants in this study were 120 first-year college students majoring in English, who were evenly divided into an experimental group and a control group, with 60 students in each [26]. The experimental group was subjected to immersive contextual English instruction, while the control group followed the conventional English teaching approach. To reduce the potential impact of sample heterogeneity on the study outcomes, the two groups were matched in terms of age, gender, and English proficiency, which was categorized based on their entrance exam scores [27]. The specific details of the experimental and control groups are presented in Table III. As indicated in the table, both groups comprised 60 participants with average ages of 18.7 and 18.8 years, respectively. A t-test revealed that the age difference between the two groups was insignificant (p=0.653). In terms of gender distribution, the experimental group had 32 males and 28 females, whereas the control group had 31 males and 29 females. A chi-square test indicated no significant difference in gender ratio between the two groups (p=0.853). The mean entrance exam scores for English were 82.3 and 81.9 for the two groups, respectively, and a t-test showed no significant difference in English proficiency (p=0.712). Regarding family background, 35 participants in the experimental group were from urban areas and 25 from rural areas; in the control group, 36 were from urban areas and 24 from rural areas. A chi-square test demonstrated no significant difference in family background between the groups (p=0.896). The duration of English study was 6.5 years and 6.4 years for the two groups, respectively, and a t-test found no significant difference in this regard (p=0.578). In summary, the experimental and control groups exhibited no significant differences in any of the aforementioned basic characteristics, suggesting that the two groups were highly comparable at the outset of the study.

TABLE III SAMPLE GROUPING AND BASIC STATISTICAL CHARACTERISTICS

Variant	Experimental Group (N=60)	Control Group (N=60)	Statistical Test Value	P-Value
Average age (years)	18.7 ± 0.9	18.8 ± 0.8	t = 0.45	0.653
Gender ratio (M/F)	32/28	31/29	$\chi^2 = 0.034$	0.853
The average score on the English test for admission	82.3 ± 5.7	81.9 ± 6.1	t = 0.37	0.712
Family background (urban/rural)	35/25	36/24	$\chi^2 = 0.017$	0.896
Length of English language study (years)	6.5 ± 1.1	6.4 ± 1.0	t = 0.56	0.578

C. Data Collection and Processing

1) Data collection: The data collection for this study covered two main types of data: quantitative and qualitative. Quantitative data were mainly obtained through standardized English test scores, which consisted of four parts: listening, reading, writing, and speaking, and were designed to comprehensively assess students' English proficiency [28]. Qualitative data collection is more diverse and includes instructional observation records, student interviews, and questionnaires. Observation records focus on student engagement and interaction in the classroom to capture the dynamics of the teaching and learning process. Student interviews were conducted to explore students' affective experiences and learning outcomes to obtain their direct feedback on teaching methods and content [29]. In addition, student satisfaction, self-confidence, and changes in interest in teaching and learning were assessed through questionnaires, the content of which helped to quantitatively analyze students' attitudes and affective responses to learning [30]. Synthesizing the data collected through multiple channels, we were able to gain a comprehensive and in-depth understanding of students' learning status and teaching effectiveness, which provided solid data support for the study.

2) Data pre-processing: The collected data were first subjected to data cleaning and pre-processing through SPSS, including outlier removal and data normalization. The data normalization equation is as follows, for the original data x, normalized to:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

Where $X_{\min}$ and $X_{\max}$ are the minimum and maximum values of the sample data, respectively?

D. Fuzzy Evaluation Model Construction

Based on the core logic of the fuzzy evaluation method, the model construction is divided into the following steps:

1) Determination of the evaluation indicator system: Combining the characteristics of the effect of immersion teaching and previous studies, an evaluation system including four first-level indicators and several second-level indicators is established [31]. The first-level indicators include: language ability (A_1), learning attitude (A_2), learning interest (A_3), and emotional experience (A_4). The secondary indicators include listening (A_{11}), speaking (A_{12}), reading and writing sub-competencies in language proficiency, classroom participation, and change in language anxiety. The weights of the first-level indicators are noted as $W = [w_1, w_2, \ldots, w_n]$, where w_i meet $\sum w_i = 1$. Each secondary indicator weight is obtained by hierarchical analysis (AHP) and distributed in the evaluation system in the following way. Table IV shows the weight distribution table of the evaluation indicators, with a total of four first-level indicators (A1, A2, A3, A4), whose sum of

weights is 1, reflecting the importance weights of different dimensions in the total evaluation. Each level 1 indicator is subdivided into several level 2 indicators, and the weight distribution is calculated by the hierarchical analysis method (AHP) to ensure rationality and scientificity [32]. The sum of the weights of the second-level indicators satisfies the corresponding weights of the first-level indicators. The distribution of the indicators reflects the multidimensional nature of the immersive contextual teaching evaluation model. For example, language proficiency is given the highest weight (40%), with higher weights for listening and speaking, indicating its centrality to teaching effectiveness.

TABLE IV DISTRIBUTION OF WEIGHTS OF EVALUATION INDICATORS

Primary Indicators	Weight	Secondary Indicators	Weight
Language Ability (A1)	WA1=0.4	Listening (Listening, A11)	W*A* 11=0.15
		Speaking (Speaking,*A* 12)	W*A* 12=0.15
		Reading (Reading,*A* 13)	WA13=0.05
		Writing (Writing,*A* 14)	WA14=0.05
Learning Attitude (A2)	WA2=0.3	Classroom Participation,*A* 21)	WA21=0.2
		Task Completion (Task Completion,*A* 22)	WA22=0.1
Learning Interest, A3)	WA3=0.2	Interest Improvement (Interest Improvement,*A* 31)	WA31=0.2
Emotional Experience (A4)	WA4=0.1	Language Anxiety Change (Language Anxiety Change,*A* 41)	WA41=0.1

2) Construct fuzzy affiliation function and affiliation matrix: Each secondary indicator corresponds to a different evaluation grade (e.g., excellent, good, fair, poor), and defines the affiliation function for each grade. Define the evaluation level set $\mathrm{V}=\{\mathrm{v}_1,v_2,...,v_m\}$, the corresponding evaluation levels include "excellent", "good", "fair", "poor", etc. For each indicator, define its fuzzy affiliation function. For each indicator A_{ij} , define its fuzzy affiliation function $\mu_{\mathrm{ij}}(x)$. In this study, triangular fuzzy numbers are used for evaluation in the following form:

$$\mu_{\mathrm{ij}}(x)\begin{cases}0, & x\le a \text{ or } x\ge c,\\ \dfrac{\mathrm{x}-a}{b-a}, & \mathrm{a}\le \mathrm{x}\le \mathrm{b},\\ \dfrac{c-x}{c-b}, & b\le x\le c\end{cases} \tag{2}$$

Where*a*, *b* and *c* are the starting point, median and end point of the fuzzy number respectively. For each evaluated object, the affiliation value of each evaluation index is calculated through the data of students' test scores, classroom records, and questionnaire feedback, and the affiliation is formed.

$$\mathrm{R}=\begin{bmatrix}\mu_{11}(v_1) & \mu_{11}(v_2) & \cdots & \mu_{11}(v_m)\\ \mu_{12}(v_1) & \mu_{12}(v_2) & \cdots & \mu_{12}(v_m)\\ \vdots & \vdots & \ddots & \\ \mu_{1n}(v_1) & \mu_{1n}(v_2) & \cdots & \mu_{1n}(v_m)\end{bmatrix} \tag{3}$$

Where each row of R represents the affiliation of a secondary indicator to a different evaluation level.

3) Fuzzy synthesis operation: Based on the affiliation matrix R and the weight vector W , a fuzzy comprehensive evaluation is performed by the following equation:

$$B=W\cdot R=\left[\mathrm{b}_1,b_2,\cdots,b_m\right] \tag{4}$$

Where B is the comprehensive affiliation vector, which represents the comprehensive evaluation affiliation of the evaluation object on different levels; W is the weight vector; and *R* is the affiliation matrix. To quantify the evaluation results, the comprehensive affiliation vector *B* is normalized and the fuzzy score is calculated:

$$\mathrm{S}=\sum_{i=1}^{m}b_i\square v_i \tag{5}$$

Where S is the final score of the fuzzy comprehensive evaluation and V_{i} is the score corresponding to the *i* evaluation level (e.g., 100 for "excellent" and 80 for "good").

E. Model Validation

To verify the applicability of the fuzzy evaluation model, this study conducted a correlation analysis between the model calculation results and the standardized test scores and further verified the reasonableness and scientificity of the model output results through expert review and student feedback. The correlation between the fuzzy composite scores and the standardized test scores was examined through Pearson correlation coefficient analysis to assess the reliability and validity of the model. The equation is as follows:

$$r = \frac{\sum(X_i - \overline{X})\sum(Y_i - \overline{Y})}{\sqrt{\sum(X_i - \overline{X})^2 \cdot \sum(Y_i - \overline{Y})^2}} \quad (6)$$

Where X_i and Y_i are the academic fuzzy score and test score, respectively, and $\overline{X}$ $\overline{Y}$ are the corresponding mean values. Based on the experimental data, the model parameters (e.g., the shape of the affiliation function and the weight distribution) are adjusted to ensure high consistency between the model evaluation results and the actual situation.

IV. Results and Discussion

Through the fuzzy evaluation of the teaching effectiveness of the experimental and control groups, this study reveals the effectiveness of immersion contextualized teaching in enhancing students' language proficiency and learning attitudes [33]. In this part, specific discussions will be made around the results of the data analysis, including the results of the fuzzy comprehensive evaluation, the analysis of the differences between the groups, and the substantive interpretation of the teaching effectiveness.

A. Fuzzy Synthesized Evaluation Results

Utilizing the previously mentioned fuzzy evaluation model, this research carried out an extensive assessment of the teaching outcomes for both the experimental and control groups, and computed their respective final fuzzy scores, as depicted in Fig. 3. This figure contrasts the comprehensive scores of the two groups. The experimental group's comprehensive fuzzy score was notably higher than that of the control group, suggesting that immersive contextual teaching can substantially enhance students' overall language abilities, particularly in communicative competencies like listening and speaking, where the benefits are more pronounced [34]. In the graph, the horizontal axis denotes the Overall Score, while the vertical axis indicates the Group, with the control group symbolized by green dots and the experimental group by purple dots. The graph reveals that the experimental group generally achieved a higher Overall Score than the control group. The score distribution in the experimental group is more clustered, with the majority of scores exceeding 80, whereas the control group's scores are more scattered, primarily ranging between 70 and 80. The experimental group's scores are more tightly packed within the 80 - 90 range, indicating greater consistency and superior performance. In contrast, the control group's scores are more diffused, spanning from 60 to 90, reflecting larger individual variations. The density curves at the base of the figure further illustrate the score distributions of both groups. The experimental group's density curve is more peaked, with the apex situated around 80 points, while the control group's density curve is flatter, with the peak around 70 points. There is a marked disparity in overall scores between the experimental and control groups, with the experimental group markedly outperforming the control group. This could imply that the pedagogical approaches or interventions employed in the experimental group were more efficacious in elevating students' overall scores.

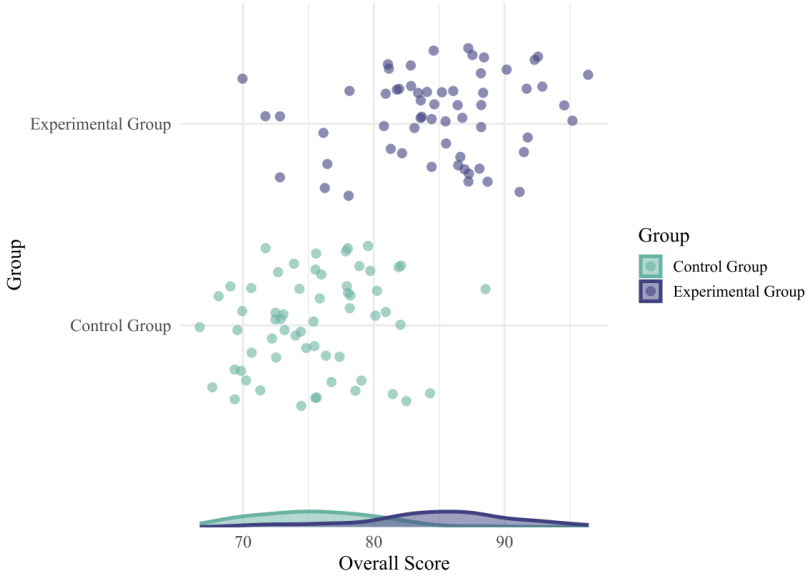

Fig. 3. Comparison of composite scores between experimental and control groups.

B. Analysis of differences between Indicators

Further analysis of the scores for the primary and secondary indicators reveals the following salient features:

1) Effective in improving language skills: The fuzzy evaluation indicates that the experimental group outperformed the control group across all language proficiency subcategories, namely listening, speaking, reading, and writing. For instance, in the listening subcategory, the experimental group's membership distribution predominantly centers on the "excellent" and "good" levels, whereas the control group's distribution is more concentrated on the "good" and "fair" levels. This suggests that immersive contextual teaching facilitates

students' quicker adaptation to real - life language settings and enhances their language comprehension abilities. Table V presents the performance membership matrices for the experimental and control groups in listening and speaking, featuring four levels of membership distribution: excellent, good, fair, and poor. Regarding listening, the experimental group's membership values for excellent and good grades are 0.45 and 0.40, respectively, which are markedly higher than those of the control group (0.10 and 0.60). In terms of speaking ability, the experimental group's membership values for excellent and good grades are 0.50 and 0.35, respectively, also surpassing those of the control group (0.15 and 0.55). This demonstrates that the experimental group exhibits superior overall performance in both listening and speaking, especially in the excellent grade category. Notably, the disparity is more pronounced in the excellent grade performance. Moreover, the membership values for the poorer grades in both groups are 0.05, indicating a low proportion of poor performance.

TABLE V LISTENING AND SPEAKING AFFILIATION MATRIX

Group	Excellent	Favorable (Good)	General (Average)	Mediocre (Poor)
Listening - Experimental Group	0.45	0.40	0.10	0.05
Listening - Control Group	0.10	0.60	0.25	0.05
Speaking - Experimental Group	0.50	0.35	0.10	0.05
Speaking - Control Group	0.15	0.55	0.25	0.05

2) Improvement of positive learning attitudes: The survey results revealed that students in the experimental group exhibited markedly higher classroom engagement, enthusiasm for learning, and associated satisfaction scores compared to the control group (refer to Fig. 4 below), which utilizes a scatter plot coupled with a trend smoothing line to illustrate a contrast between the experimental and control groups regarding learning attitude scores. As depicted in the figure, the experimental group generally achieved higher learning attitude scores than the control group. Despite both groups having a certain degree of score distribution, the experimental group's scores were more concentrated and displayed an upward trend, indicating that students in this group demonstrated a more proactive and positive approach to learning [35]. Conversely, the control group's scores were relatively lower and more scattered, suggesting that their learning attitudes might be less favorable than those of the experimental group. Moreover, the experimental group had a significantly greater number of individuals with high ratings than the control group, further corroborating the efficacy of the experimental teaching method in bolstering students' learning attitudes [36]. In summary, the figure furnishes visual substantiation for the study that the teaching interventions in the experimental group yielded substantial outcomes in enhancing students' attitudes towards learning. This underscores the efficacy of contextualized teaching in stimulating students' intrinsic motivation.

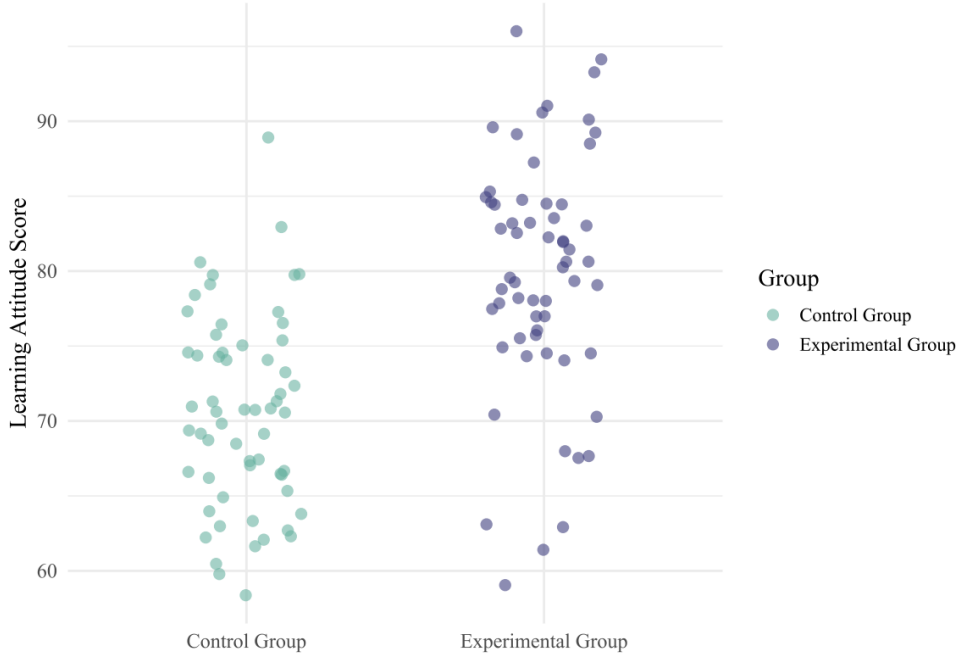

Fig. 4. Statistical comparison of learning attitude indicators between groups.

C. Comparison of Teaching Effectiveness between Experimental and Control Groups

To present a more comprehensive picture of the differences in the effectiveness of immersive contextualized instruction versus traditional instruction, this study conducted a t-test on the composite scores of the two samples, yielding in the Fig. 5.

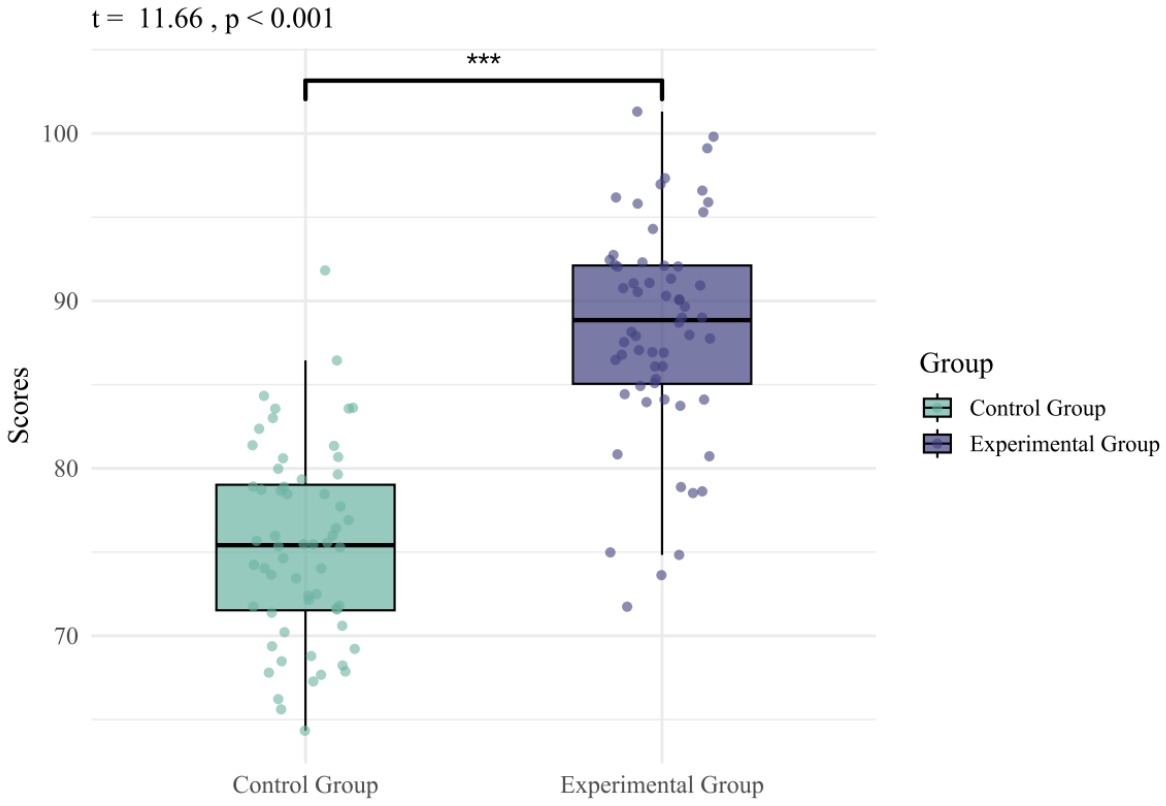

Fig. 5. Graph comparing the effect of experimental and control groups.

The figure provides a detailed comparison of teaching effectiveness between the experimental and control groups, depicted through a box-and-line plot. It is evident that the experimental group's score distribution is more compact and features a higher median, suggesting superior overall performance compared to the control group. Specifically, the experimental group's median score is nearly 90, whereas the control group's median score hovers around 75, highlighting a distinct advantage in teaching effectiveness for the experimental group. Furthermore, the experimental group's score range is relatively narrower, indicating more stable and less variable student performance. The statistical analysis results corroborate this observation. The independent samples *t*-test results revealed a highly significant difference between the two groups (t=11.66, p<0.001), indicating that the experimental group was substantially more effective than the control group. This level of significance (p<0.001) implies that the difference between the experimental and control groups is improbable to be attributed to random factors, but rather to the efficacy of the teaching methods or interventions employed by the experimental group.

D. Interpretation and Discussion of Results

1) Advantages of immersive contextualized instruction: In this study, the effects of immersive contextual teaching were analyzed in depth through the fuzzy evaluation method, and the results showed that this teaching mode can effectively improve students' language communication skills. This finding is consistent with the existing literature on the importance of authenticity of teaching contexts for students' language acquisition [37]. What's more, this study is the first attempt to quantitatively analyze teaching effectiveness through the fuzzy evaluation method, which provides a new perspective and methodology for research in this field. The results of this quantitative analysis not only enhance the scientific nature of teaching research but also provide precise data support for actual teaching design, which helps educators grasp the teaching effect more accurately to optimize teaching strategies.

2) Application value of fuzzy evaluation: Fuzzy evaluation shows its unique advantages in assessing complex teaching phenomena, especially in the multidimensional comprehensive analysis [38]. For example, when assessing highly subjective indicators such as "emotional experience", the affiliation matrix of fuzzy evaluation can effectively describe the distribution of student's satisfaction with the classroom, thus reducing the bias that may be brought by a single scoring model [39]. This method can reflect students' subjective feelings and learning experiences more comprehensively, providing a richer and more detailed perspective for the evaluation of teaching effectiveness.

3) Research limitations and future directions: While this study has yielded significant findings regarding the assessment of teaching effectiveness, it is not without limitations. Firstly, the relatively small sample size may impinge upon the generalizability of the results. Subsequent research endeavors could enhance the generalizability and robustness of the model presented herein by increasing the sample size. Secondly, the parameter configurations of the fuzzy evaluation model, including the form of the membership function, warrant further refinement to augment the model's precision and versatility [40]. Moreover, future investigations might delve into the applicability of fuzzy evaluation techniques across diverse teaching contexts and subject domains, thereby broadening the scope and depth of their utilization. Through such initiatives, it is anticipated that fuzzy evaluation methods will assume a more pivotal role in the realm of teaching effectiveness assessment, offering enhanced support for educational research and practice.

V. Conclusion and Limitations

This study adopts a fuzzy evaluation method to comprehensively analyze the effectiveness of immersive contextual English teaching. By constructing a scientific fuzzy evaluation model, this study assessed the effectiveness of the teaching method in terms of two key dimensions, namely, language proficiency and learning attitude, and proved the significant advantages of immersive contextual teaching in a practical application through empirical research. The results of the study show that immersive contextual teaching is effective in enhancing students' language proficiency. The composite scores of students in the experimental group were significantly higher than those of the control group in each language proficiency index, such as listening, speaking, reading, and writing. This result shows that contextualized teaching can effectively improve students' language practice ability by creating a real language environment. In addition, immersion teaching also shows positive effects in enhancing students' learning attitudes. The results of the fuzzy evaluation show that immersion teaching is effective in enhancing students' motivation, classroom participation, and interest in learning. Scenario simulation and interactive experience can better stimulate students' intrinsic motivation and thus strengthen learning effects. This study also introduces the fuzzy evaluation method into the assessment of English teaching effectiveness and verifies the application value of this method. This method provides a new way of thinking and methodology for the teaching evaluation system, which can quantify the complex teaching process more comprehensively, especially when it involves the comprehensive analysis of multi-dimensional data, which has significant advantages. To summarize, immersive contextual English teaching not only reflects high efficiency in language proficiency cultivation but also achieves positive feedback in students' learning attitudes and subjective experiences, providing important insights for educational reform and practice. The application of the fuzzy evaluation method further highlights the importance of quantitative analysis in teaching research, which is of theoretical promotion and practical guidance significance.

However, there are still some shortcomings in this study. On the one hand, the relatively small size of the experimental sample may affect the external validity of the conclusions; future research can expand the sample coverage and select subjects from students of different age groups and different language bases to verify the generalizability of the findings. On the other hand, the fuzzy evaluation model is somewhat subjective in the selection of the affiliation function; in the future, attempts can be made to optimize the parameter settings of the model by introducing machine learning or artificial intelligence algorithms. In addition, this study mainly focuses on the two dimensions of language proficiency and learning attitude, and subsequent studies can further explore the effects of contextualized teaching on higher-order language proficiency such as critical thinking and intercultural communication skills.

References

[1] Al-Gerafi M, Goswami S, Khan M. Designing of an effective e-learning website using inter-valued fuzzy hybrid MCDM concept: A pedagogical approach[J]. *Alexandria Engineering Journal*, 2024, 97: 61–87.

[2] Jiang L. Research on the integration path and practice of AI intelligent technology and English distance education[J]. *Sciendo*, 2023, 45(6): 23–50. doi: 10.2478/amns.2023.2.01427.

[3] Al-kfairy M, Ahmed S, Khalil A. Factors impacting users' willingness to adopt and utilize the metaverse in education: a systematic review[J]. *Comput Hum Behav Rep*, 2024, 76(56): 100459.

[4] Al-Samarraay M, Salih M, Ahmed M. A new extension of FDOSM based on Pythagorean fuzzy environment for evaluating and benchmarking sign language recognition systems[J]. *Neural Comput Appl*, 2022, 56(5): 1–19.

[5] Caiado R, Scavarda L, Gavião L. A fuzzy rule-based industry 4.0 maturity model for operations and supply chain management[J]. *Int J Prod Econ*, 2021, 231: 107883. doi: 10.1016/j.ijpe.2020.107883.

[6] Mikhailenko M, Maksimenko N, Kurushkin M. Frontiers | eye-tracking in immersive virtual reality for education: a review of the current progress and applications[J]. *Sciendo*, 2023, 34(4): 13–25. doi: 10.3389/feduc.2022.697032.

[7] Cheng H, Zhu L, Meng J. Fuzzy evaluation of the ecological security of land resources in mainland China based on the pressure-state-response framework[J]. *Sci Total Environ*, 2022, 804: 150053.

[8] Chien C, Ho Y, Hou H. Integrating immersive scenes and interactive contextual clue scaffolding into a decision-making analysis ability training game[J]. *J Educ Comput Res*, 2024, 62(1): 376–405.

[9] Purnama Y, Fransiska F, Muhdi A. Long-term effects evaluation of using artificial intelligence-based automated learning systems in improving English content understanding at the secondary education level[J]. *Indones J Educ (INJOE)*, 2023, 3(3): 622–636.

[10] Chua C, Kosnin A, Yeo K. Fuzzy Delphi method for a-level mathematics technological pedagogical and content knowledge module[J]. *Int J Eval Res Educ*, 2024, 13(1): 441–453.

[11] Yuan R. The other side of the coin: A socio-cultural analysis of pre-service language teachers' learning to teach critical thinking[J]. *Think Skills Creat*, 2023, 48: 101265. doi: 10.1016/j.tsc.2023.101265.

[12] Restall G, Yao Y, Niu X. Exploring the experience of year 10 South Korean students' English language learning in immersive virtual reality[J]. *TESOL Context*, 2023, 31(2): 21–67.

[13] Yuan R. Cultivating CT-oriented teachers in pre-service teacher education: what is there and what is missing?[J]. *Taylor Fr*, 2023, 76(9): 67–99.

[14] Sanfilippo F, Blazauskas T, Salvietti G. A perspective review on integrating VR/AR with haptics into STEM education for multi-sensory learning[J]. *Robotics*, 2022, 11(2): 41. doi: 10.3390/robotics11020041.

[15] Meccawy M, Alzahrani A, Mattar Z. Assessing EFL students' performance and self-efficacy using a game-based learning approach[J]. *Educ Sci*, 2023, 13(12): 1228. doi: 10.3390/educsci13121228.

[16] Curran V, Xu X, Aydin M. Use of extended reality in medical education: an integrative review[J]. *Med Sci Educ*, 2022, 33(1): 275–286. doi: 10.1007/s40670-022-01698-4.

[17] Yuchen X. Application of immersive artificial intelligence based on machine vision in education management of children with autism[J]. *Int J Syst Assur Eng Manag*, 2023, 28(4): 1–10.

[18] Fang C. Intelligent online English teaching system based on SVM algorithm and complex network[J]. *J Intell Fuzzy Syst*, 2021, 40(2): 2709–2719.

[19] Shi L, Muhammad Umer A, Shi Y. Utilizing AI models to optimize blended teaching effectiveness in college-level English education[J]. *Cogent Education*, 2023, 10(2): 2282804.

[20] Manabe K, Hwang W, Chuang Y. English learning is enhanced by collaborative contextual drama in an authentic context[J]. *Interact Learn Envir*, 2023, 31(7): 4490–4506.

[21] Zhou S. Gamifying language education: the impact of digital game-based learning on Chinese EFL learners[J]. *Humanit Soc Sci Commun*, 2024, 11(1): 1–14. doi: 10.1057/s41599-024-04073-3.

[22] Lu C, He B, Zhang R. Evaluation of English interpretation teaching quality based on GA optimized RBF neural network[J]. *J Intell Fuzzy Syst*, 2021, 40(2): 3185–3192.

[23] Li Y. The digital transformation of college English classroom: application of artificial intelligence and data science[J]. *EAI Endorsed Trans Scalable Inf Syst*, 2024, 11(5): 23–53.

[24] Song C, Shin S-Y, Shin K-S. Optimizing foreign language learning in virtual reality: a comprehensive theoretical framework based on constructivism and cognitive load theory (VR-CCL)[J]. *Appl Sci*, 2023, 13(23): 12557.

[25] Zhang L. An IoT-based English translation and teaching using particle swarm optimization and neural network algorithm[J]. *Soft Comput*, 2023, 27(19): 14431–14450.

[26] Zhang Y. An analytical study on the design and implementation of a comprehensive web-based learning environment for Chinese as a foreign language[J]. *Soft Computing*, 2023, 27(23): 18147–18164.

[27] Li N. A fuzzy evaluation model of college English teaching quality based on analytic hierarchy process[J]. *Int J Emerg Technol Learn (iJET)*, 2021, 16(2): 17–30.

[28] Khalaf OI, Srinivasan D, Algburi S. Elevating metaverse virtual reality experiences through network-integrated neuro-fuzzy emotion recognition and adaptive content generation algorithms[J]. *Eng Rep*, 2024, 34(10): e12894.

[29] Wang H, Wang J, Wang G. A survey of fuzzy clustering validity evaluation methods[J]. *Inf Sci*, 2022, 618: 270–297.

[30] Zolghadri M, Jafarpour Mamaghani H. Pathology of language teaching ineffectiveness: a case study exposing teacher cognition stepwise[J]. *J Mod Res Engl Lang Stud*, 2021, 9(1): 29–51.

[31] Jin S, Huang J, Zhong Z. Application of immersive technologies in primary and secondary education[J]. *Front Digit Educ*, 2024, 1(2): 142–152.

[32] Huang J, Sang G. Conceptualising critical thinking and its research in teacher education: a systematic review[J]. *Teach Teach*, 2023, 29(6): 638–660.

[33] He Q, Attan S, Zhang J. Evaluating music education interventions for mental health in Chinese university students: a dual fuzzy analytic method[J]. *Sci Rep*, 2024, 14(1): 19727.

[34] Wang W, Huang S. The application of artificial intelligence teaching software in college English teaching[J]. *Sciendo*, 2021, 34(34): 44–62. doi: 10.2478/amns.2023.2.00657.

[35] Han L. Students' daily English situational teaching based on virtual reality technology[J]. *Mobile Inf Syst*, 2022, 2022(1): 1222501.

[36] Zheng B. Translanguaging in a Chinese immersion classroom: an ecological examination of instructional discourses[J]. *Int J Biling Educ Bi*, 2021, 133(34): 126–144.

[37] Xie Y, Liu Y, Zhang F. Virtual reality-integrated immersion-based teaching to English language learning outcome[J]. *Front Psychol*, 2022, 12: 767363.

[38] Weng Y, Schmidt M, Huang W. The effectiveness of immersive learning technologies in K–12 English as second language learning: a systematic review[J]. *Recall*, 2024, 45(6): 1–20.

[39] Gong J, Liu H, You X. An integrated multi-criteria decision-making approach with linguistic hesitant fuzzy sets for E-learning website evaluation and selection[J]. *Appl Soft Comput*, 2021, 102: 107118.

[40] Fernández-Herrero J. Evaluating recent advances in affective intelligent tutoring systems: A scoping review of educational impacts and prospects[J]. *Humanities and Social Sciences Communications*, 2024, 122(16): 67–98. doi: 10.3390/educsci14080839.