Journal of Combinatorial Mathematics
and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Measuring the Effectiveness of Students' Language Proficiency Enhancement Based on Principal Component Analysis in Task-Based Teaching of English in Information Technology-Assisted Colleges and Universities

Peilan Peng[1], Yuanlu Meng[2,] ✉

[1] *ZHIYUAN (Liberal Study) College, Beijing Institute of Petrochemical Technology University, Beijing, 102617, China*

[2] *College of New Materials and Chemical Engineering, Beijing Institute of Petrochemical Technology University, Beijing, 102617, China*

ABSTRACT

The traditional English teaching mode in colleges and universities has many problems in cultivating students' language ability. This paper introduces information technology into task-based English teaching in colleges and universities and constructs a task-based English teaching mode based on SPOC technology. With the orientation of improving students' language ability, it implements the improvement of English teaching mode in colleges and universities. Using principal component analysis to comprehensively evaluate the relevant indicators of students' language proficiency in the process of task-based English teaching in colleges and universities, and quantify the effect of the combination of information technology and task-based English teaching on the improvement of students' language proficiency. Ten classes of students majoring in English in a university were selected and divided into experimental and control groups, and the data related to students' language proficiency were collected and analyzed at the end of the experiment. The data were downscaled using principal component analysis, and the principal components were extracted according to the eigenvalues and cumulative contribution rate. The comprehensive score of students' language proficiency is calculated by the comprehensive evaluation function of students' language proficiency constructed in this paper. The language proficiency of students in the experimental group and the control group is significantly different after the experiment, and the comprehensive scores of students in the experimental group are 53.96% and 61.96% higher than those before the experiment, respectively. It reveals that the introduction of information technology into task-based teaching of English in colleges and universities has a significant effect on the enhancement of students' language proficiency.

✉ Corresponding author.
E-mail address: lansunny1115@126.com (Y. Meng).
Received 25 February 2024; Revised 10 May 2024; Accepted 15 December 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-235

Keywords: information technology, task-based teaching, principal component analysis, cumulative contribution rate, students' language proficiency

1. Introduction

Task-based language teaching is to complete specific tasks as the learning motivation and motive, the process of completing tasks as the learning process, and to demonstrate the achievement of the task results to show the achievement, which is different from the traditional one-way lecture mode teaching [1-2].

The core idea of task-based language teaching is to simulate all kinds of activities that people engage in by using language in the real life of society and school, and to combine language teaching with learners' language application in daily life [3-4]. And English task-based teaching is a task-centered approach to English teaching, which is a way for teachers and students to let students complete a series of teaching tasks designed according to their developmental needs through English dialogue, communication, etc., so that students can achieve their learning goals and realize cross-cultural communication and innovation by doing things in English [5-6]. English task-based teaching emphasizes that based on a variety of English learning tasks, students must have a thinking process in the process of completing the tasks, i.e., students should first consider how to complete the learning tasks, rather than how to learn the English language form, which transforms the basic concept of English language application into a classroom teaching method with practical significance. English task-based teaching pays attention to the process of English learning, emphasizes the multilateral interaction between students and between students and teachers, and tries to create a natural and authentic native English environment, so that students can use English through negotiation and communication of meanings and concrete events in the process of completing the tasks, thus developing students' English language proficiency, especially their English communicative competence [7-9]. The purpose of English task-based teaching is to promote the improvement of learners' English proficiency and to cultivate students' abilities of observation, thinking, memorization, imagination and innovation, so as to improve the overall quality of students [10].

The implementation of English task-based teaching cannot be separated from the support of information technology. With the network and multimedia technology as the core of modern information technology can optimize or change the teaching environment, make the teaching environment more flexible, so that the teaching designers and learners enjoy a more suitable teaching environment, so as to meet a variety of teaching needs. In the implementation process of task-based teaching, information technology is used as a variety of tools to improve learning efficiency, develop thinking ability, stimulate enthusiasm for learning, and enhance students' cognitive efficiency and ability [11-13].

Now more and more colleges and universities are paying obvious attention to the improvement of students' English language proficiency, facilitating the development of students' oral skills for better communication with foreign students and the improvement of their own second language literacy. Task-based English teaching has gradually become the norm in the English classroom and the necessity of implementation in the current form [14], and students' interest in learning the English language and knowledge and understanding of the English language has increased [15]. And with the support of information technology, students' English language proficiency improves [16]. When information technology is combined with task-based English teaching, students' English language proficiency is significantly improved, on the one hand, the interactivity of task-based teaching strengthens communication, on the other hand, information technology makes English knowledge more visual and three-dimensional strengthens students' comprehension [17]. And virtual reality technology is very common immersive, interactive, intuitive and effective information assistive

technology in English language learning [18]. However, the improvement of students' English language skills is not clear, which is very unfavorable for targeted learning. However, there are many factors affecting students' English language proficiency in information technology-assisted task-based teaching, for which principal component analysis can realize the representativeness of the constraints and the close connection between the factors to extract the core variables [19]. Literature [20] then used the results of principal component analysis to show that gamified task-based teaching has a positive effect on student achievement.

This paper takes SPOC technology as a representative of information technology and introduces it into task-based English teaching in colleges and universities to improve the traditional college English teaching mode in order to improve students' language proficiency. The data of English majors in a selected college after the experiment are analyzed by using principal component analysis. By comparing the differences in the results of the comprehensive scores of language proficiency between the experimental group and the control group, and the changes in the comprehensive scores of language proficiency of the experimental group before and after the experiment, we highlight the effect of utilizing information technology to assist task-based English teaching in colleges and universities in improving the language proficiency of the students.

2. Information Technology Aids Task-Based Teaching of English in Colleges and Universities

2.1. Task-based English Teaching

Task-based English teaching has three very important phases [21], i.e., the pre-task phase, the task chain phase, and the linguistic analysis phase, which are specifically included as follows.

1) Pre-task stage. In this stage, some activities are needed to introduce the content to be learned, through which students are stimulated to recall the content and background of what they have learned as well as the information related to the knowledge in their minds to help them rebuild the language knowledge system. At the same time, the students can be fully equipped with the ability to learn the current stage of language knowledge and cultural knowledge, which will reduce the pressure on the next stage of language knowledge learning and improve the students' active ability to learn language knowledge.

2) Task Chain Phase. This stage is mainly to do the language learning tasks, prepare the language learning report, report the results of the language learning tasks, as the language learning stage, teachers need to pay attention to the fluency of the students in English language and the accuracy of the mastery of English language learning in the process, choose the appropriate difficulty of the task in a timely manner to alleviate the pressure of the students in the process of learning, and encourage students to rebuild the enthusiasm for English learning as well as specific strategies.

3) Language Analysis Phase. In this stage, students need to independently demonstrate the process of accomplishing the language learning tasks, through which they can find out their own shortcomings and areas for improvement, and establish a set of their own clear-cut language forms and functions. Teachers emphasize and practice specific language forms and provide feedback on the students' performance in the previous stage of language learning. This staged presentation and feedback provides an opportunity for students to re-learn and reflect on the tasks they have done in the previous stage, and to gain a better understanding of the language forms.

2.2. The Practice of Information Technology in Task-Based English Language Teaching

SPOC is one of the efficient learning modes brought about by the current development of information technology, which makes up for the shortcomings of traditional task-based English learning in colleges and universities [22]. This section applies SPOC technology to task-based English teaching in colleges and universities to realize the integration of information technology and task-based English teaching.

2.2.1. Objects of implementation:

1) Teachers. Teachers are the main participants and pedagogues of teaching activities, occupying an important role in the whole teaching process, so teachers in the teaching mode need to have several conditions: First, teachers should have basic information technology literacy and be able to master computer network technology. Secondly, they should be extremely familiar with the whole process of online network learning and have a certain understanding of the various learning APPs of MOOC. Third, they should abandon the traditional teaching methods and focus on communicating with students and guiding learning for learning.

2) Students. Students are the main participants and recipients of teaching, strictly speaking, the learning effect depends on the level of the recipient, and only the recipient can fully absorb the knowledge by taking the initiative to learn. Therefore, students also need to have certain literacy. First, they should master basic English vocabulary and have a certain level of grammar. Secondly, they should have the awareness of utilizing online resources for learning and be able to operate smart phones, computers and other devices.

2.2.2. Modes of instruction. This paper constructs a task-based English teaching model for colleges and universities based on SPOC from five aspects: teaching objectives, teaching objects, teaching strategies, teaching process and teaching evaluation.

Figure 1 shows the framework of task-based English teaching model in colleges and universities based on SPOC, which is divided into pre-teaching, teaching and post-teaching, with five modules running through the whole teaching process.

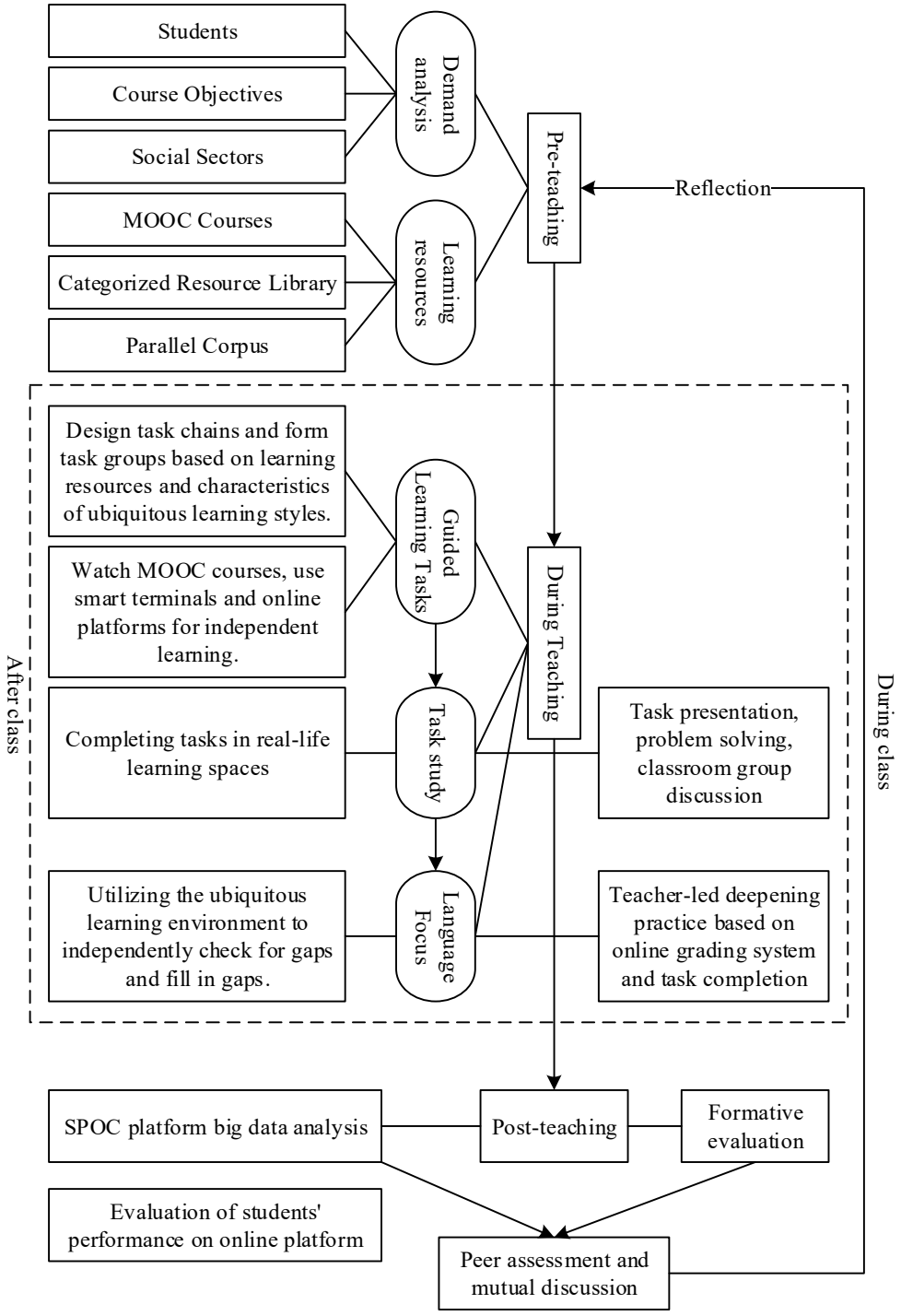

Fig. 1. The college task type English teaching model based on SPOC

1) Teaching objectives. Teaching objectives is an important factor to be determined before teaching, this link needs to analyze the needs of students, the actual needs of students, the social demand for human skills and the purpose of the course needs to be achieved as an important consideration, the trinity, the formation of teaching objectives. The teaching objectives of this teaching mode is the main consideration is the needs of students, focusing on the development of student autonomy, and strive to promote the overall development of students through perfect teaching.

2) Teaching Objects. The analysis of the teaching object needs to start from two aspects, the first is to understand the cognitive level of students before the start of teaching activities, and then sort out the relevant teaching resources and match the corresponding learning resources according to the different

levels of students. The learning resources are mainly MOOC courses, categorized resource databases, and parallel corpora.

3) Teaching Strategies. Teaching strategy refers to a series of teaching methods adopted in teaching activities in order to accomplish the established teaching goals. English language task-based teaching mode of SPOC is a teaching mode designed based on the traditional SPOC teaching methods. Students learn independently based on the Internet teaching resources, and make offline learning notes and learning questions during the period, and then the teacher explains the key learning content in classroom teaching and answers students' doubts, so as to promote the learning of English language. and answer students' questions, thus promoting the quality of learning.

4) Teaching process. The teaching process is mainly divided into task-guided learning, task study and language focus. Task-guided learning is based on the learning content set in the learning objectives, providing learning resources to guide students to learn, mainly using Internet devices, based on the MOOC course to learn, and then use classroom teaching to conduct on-site task research, which is the core link, and finally the language focus, which is to make full use of the learning environment to check the gaps and make up for the deficiencies.

5) Teaching Evaluation. This link is to evaluate the learning effect of students and the teaching effect of teachers after teaching, which are the mutual evaluation among students, the learning situation of network learning platform and the big data analysis of SPOC platform, among which the mutual evaluation of students and the evaluation of network learning platform are carried out after each course, and the big data analysis of SPOC platform is evaluated after the completion of the whole learning objectives.

2.2.3. Instructional design. As mentioned before, task-based English teaching is divided into three stages: pre-task, task chain and language analysis. This section takes the three steps of task-based teaching as a framework and combines them with specific teaching cases to explore task-based English teaching methods in colleges and universities supported by SPOC information technology. Take the teaching of Section A Five Famous Symbols of American Culture in the fourth unit of the third book of New Vision College English textbook as an example.

1) Pre-task stage. In the pre-task stage, the teacher introduces the teaching objectives and tasks to the students, which mainly include: providing authentic tasks, introducing the significance of the tasks, encouraging the students to actively participate in the implementation of the tasks, providing authentic materials in the target language, and guiding the students to pay attention to the linguistic forms by analyzing the texts and data.

Teaching Objectives:

(1) Grasp the main idea and structure of the text, improve students' listening, speaking, reading and writing abilities.

(2) Shed light on the five famous symbols of American culture and arouse students' interest in different cultures.

(3) Train students' ability to search information on the MOOC and show the result of the research.

(4) Cultivate students' ability to cooperate with others.

(5) Show students several well-known pictures of symbols of different cultures with the multimedia computer to attract the students' attention to the topic of the text and then ask students to identify these symbols in English.

(6) Ask students to discuss in groups and the answer the following questions:

What do you think are the symbols of Chinese culture?

Can you name at least one and tell something about it?

After students finish the tasks, the teacher shows the answers on the screen and explain some new words and phrases which are difficult to understand.

By completing the tasks in this stage, not only should the students master the content of the text and

the new vocabulary involved through independent study, but also make them master other knowledge related to the topic of the text through inquiry and consultation with their classmates in order to expand their knowledge.

2) Task chain stage. Task chain consists of three stages: task, plan and report. Students can learn the content of the text independently through MOOC. The teacher has to design tasks according to the topics and materials provided in the text, and students carry out activities in groups, and the teacher has to encourage students to communicate in the target language and help them express themselves systematically. Students take the initiative to speak, cooperate with each other and prepare for the report that follows. Finally, the group chooses a representative to present the results to the class and share the information. Students can accomplish this stage in various forms, such as presentation, roleplay, debate and so on, in order to further deepen their understanding of the text. At this time, the teacher should listen carefully, then summarize and give reasonable evaluation, while letting the students carry out self-evaluation and mutual evaluation.

(1) Show students pictures of the five famous symbols of American culture mentioned in the text on the screen. Then ask students to work in pairs to find out the name, time, designer, model and history of the five symbols respectively.

(2) Ask students to discuss the following question in groups: What more can you say about the symbols of American culture?

(3) Ask students to make a PowerPoint presentation.

Ten culture symbols are listed on the screen: the Statue of Liberty, Barbie, American Gothic, the Buffalo, Nickel, the Great Wall, Dragon, traditional Chinese medicine, Chinese Kongfu, and Calligraphy. Each group is required to make a PowerPoint presentation on one of the ten topics.

At this stage, students should make full use of MOOC resources to find the materials they need under the guidance of teachers, and utilize the support means of SPOC information technology to complete the above tasks. In the process of cooperative learning, with the help of the teaching environment provided by information technology, students are in a state of active learning and acquire language in the process of accomplishing the tasks.

3) Language analysis stage.After students have completed the tasks, teachers should design various forms of activities to analyze and emphasize certain language points, so that students can further consolidate the language knowledge points of the lesson. Teachers can organize some targeted exercises. In the process of completing the task, students' attention is focused on the meaning rather than the form of the language, but the teacher can balance the two by directing students' attention to the form through post-task activities.

(1) The teacher focus on the language again and explains the grammar and asks the students to practice the structures on Page 102 and 103.

(2) Students are required to discuss the question in class.

How to promote cultural exchanges between China and the United States?

3. Principal Component Analysis

Through the construction of task-based English teaching model in colleges and universities based on SPOC technology and the design of task-based teaching practice, this paper successfully introduces information technology into task-based English teaching in colleges and universities. In the next step, this paper will study the effect of introducing information technology in task-based English teaching in colleges and universities on the improvement of students' language proficiency. As one of the most commonly used data dimensionality reduction algorithms, principal component analysis can also be regarded as a multivariate statistical analysis method to grasp the main contradictions of things, and it is the most commonly used feature extraction method. Principal component analysis can be used to comprehensively evaluate the language proficiency-related indicators of students in task-based teaching of English in colleges and universities in order to measure the effect of students' language

proficiency enhancement. Therefore, this section will focus on the principal component analysis method.

3.1. Mathematical model for principal component analysis

Principal component analysis aims to use the idea of dimensionality reduction to concentrate the information dispersed on a set of variables on a few composite variables (linear combinations), and the resulting composite variables are not correlated with each other [23]. From a geometric point of view, principal component analysis is a coordinate rotation of the original axes to obtain mutually orthogonal axes, so that the direction of this axis is the direction in which all the data points are most widely dispersed, and these new axes are arranged according to the magnitude of the eigenvalues obtained.

In the new coordinate system obtained, the correlation of the data between the dimensions is minimized, in other words, this transformation is a de-correlation process, i.e., the new variables obtained are uncorrelated with each other.

The principal component analysis is described in mathematical terms: set up in data set X with n samples and P variables observed in each sample, there are:

$$X=\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}=\left[x_1,x_2,\cdots x_p\right] \tag{1}$$

Among them:

$$x_i=\left(x_{1i},x_{2i},\cdots x_{nt}\right)^T,i=1,2,\cdots p \tag{2}$$

PCA is the synthesis of the original P observed variables $x_1,x_2,\cdots x_p$ to form P new variables (composite variables), i.e:

$$\begin{cases} F_1=w_{11}x_1+w_{21}x_2+\cdots+w_{p1}x_p \\ F_2=w_{12}x_1+w_{22}x_2+\cdots+w_{p2}x_p \\ \cdots \\ F_p=w_{1p}x_1+w_{2p}x_2+\cdots+w_{pp}x_p \end{cases} \tag{3}$$

Abbreviated as:

$$F_i=w_{11}x_1+w_{21}x_2+\cdots+w_{pi}x_p,i=1,2,\cdots,p \tag{4}$$

Here, x_i is a n-dimensional vector, yielding F_i which is also a n-dimensional vector. The coefficients w_y of the above model are subject to the following three conditions:

1) F_i is not correlated with $F_j(i\neq j;i,j=1,2,\cdots p)$.

2) The variance of F_1 is greater than the variance of F_2 is greater than the variance of F_3, and so on.

3) ${w_{k1}}^2+{w_{k2}}^2+\cdots+{w_{kp}}^2=1$, $k=1,2,\cdots,p$.

When the above three conditions are satisfied, the new random variables obtained from the transformation are uncorrelated with each other and their variances decrease in descending order.

In summary, the transformation matrix of order $p\times p$ is obtained as W:

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1p} \\ w_{21} & w_{22} & \cdots & w_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ w_{p1} & w_{p2} & \cdots & w_{pp} \end{bmatrix} \tag{5}$$

Make:

$$F = \left[F_1, F_2, \cdots F_p \right] = W^T X \tag{6}$$

3.2. Principal Component Contributions

The contribution ratio and cumulative contribution ratio of principal components measure how much information the transformed F has extracted from the original data X.

1) Contribution ratio: the proportion of the eigenvalues corresponding to the i rd principal component in the sum of all eigenvalues of the covariance matrix, the larger the ratio, the stronger the ability of the i th principal component to synthesize the information of the original index. The eigenvalue corresponding to the i th principal component is λ. The formula is:

$$\alpha_i = \lambda_i \Big/ \sum_{i=1}^{p} \lambda_i \tag{7}$$

2) Cumulative contribution ratio: the proportion of the sum of the eigenvalues of the first k principal component in the sum of all the eigenvalues, the larger this ratio is, the more the first k principal components can represent the information of the original data comprehensively. The formula is:

$$M_k = \sum_{i=1}^{k} \lambda_i \Big/ \sum_{i=1}^{p} \lambda_i \tag{8}$$

In solving practical problems, generally select the first $d(d < p)$ principal component, so that its cumulative variance contribution rate to meet certain requirements (usually more than 85%), with the selection of the first d principal components instead of the original p variables to analyze, you can achieve the purpose of the data dimensionality reduction can also be regarded as a feature extraction.

3.3. Application of Principal Component Analysis in English Teaching and Learning

Principal component analysis has wide application value in the field of educational evaluation. In the evaluation of students' comprehensive quality, it often involves multiple evaluation indicators, and there may be a certain correlation between these indicators. Through principal component analysis, multiple related indicators can be transformed into a few irrelevant principal components, avoiding the problem of overlapping information and inaccurate evaluation results caused by the correlation of indicators in traditional evaluation methods. For example, when evaluating students' academic performance, which may involve multiple subject grades, regular homework grades, test scores, classroom performance and other indicators, principal component analysis can be utilized to extract the principal components that reflect the main aspects of students' abilities, learning attitudes, and learning effects, so as to more comprehensively and objectively evaluate the students' academic level. In this study, principal component analysis will be utilized to comprehensively process multiple indicators related to students' language proficiency generated from task-based teaching of English in colleges and universities, and to construct a model of the effect of students' language proficiency improvement.

4. Model for Measuring the Effectiveness of Students' Language Proficiency Improvement

4.1. Establishment of a system of measurement indicators

Based on the principles of comprehensiveness, representativeness and operability, this paper constructs the evaluation index system for the effect of students' language proficiency enhancement from the four dimensions of language knowledge, language skills, learning attitudes and strategies, and the quality of task completion, and the specific results are shown in Table 1. The dimension of language knowledge includes vocabulary and grammatical knowledge mastery. Vocabulary is measured by vocabulary testing software, and the degree of language knowledge mastery is assessed according to the language transformation test scores. The language skills indicator covers four dimensions: listening, speaking, reading and writing. Listening skills are reflected by listening comprehension test scores, such as TOEFL and IELTS listening section scores. Speaking skills can be measured by speaking interview scores, speaking competition scores, etc. Reading skills are measured by reading comprehension test scores, such as scores on the reading section of College English Level 4 or 6. Writing skills, on the other hand, are assessed based on the scoring results of writing tasks, such as course papers and essay test writing. Learning attitude and strategy dimensions included study time commitment, study plan development and execution, classroom participation, and resource utilization. Learning time is measured by students' learning log records and MOOC online learning hours statistics. Learning plan development and implementation are judged by the learning plans submitted by students and the teacher's observation and evaluation of their implementation to determine students' learning planning ability and self-management ability. The degree of students' participation in classroom interaction is quantified according to the number of classroom presentations and the enthusiasm and contribution of group discussion participation. Resource utilization ability is used to assess students' utilization of English resources provided by information technology, such as the completion rate of MOOC course learning and the frequency of using MOOC software. Task completion quality indicators contain task accuracy, task completeness and task innovativeness. Task accuracy is measured by the accuracy of language expression and the accuracy of information delivery. Task completeness is measured by whether students cover all the required content and achieve the basic objectives set for the task. Task innovativeness was measured by whether students presented new ideas, methods, and creativity in the task.

Table 1. Student language ability improvement evaluation system

Primary indicator	Secondary indicator
Language knowledge	Vocabulary(X1)
	Grammar knowledge(X2)
Linguistic ability	Listening skill(X3)
	Oral skill(X4)
	Reading skill(X5)
	Writing skill(X6)
Learning attitude and strategy	Study time input(X7)
	Study planning(X8)
	Classroom participation degree(X9)
	Resource using ability(X10)
Task completion quality	Task accuracy(X11)
	Task integrity(X12)
	Task innovation(X13)

4.2. Data collection and organization

4.2.1. Classroom observation. The main way of obtaining classroom observation data for the study was conducted in the form of video recording (videotaping and synchronized audio recording). The researcher established a good relationship and mutual trust through face-to-face contact, formal and informal communication, internet interviews, focus groups, interviews, and other means of communication and interaction with the research participants over a period of time. The researcher indicated to the research subjects during the screening stage that they needed to enter the teaching site and record the teaching process in the form of video recording; based on this, the researcher again formally indicated to the identified research subjects the above research request and the related confidentiality and anonymity measures. After obtaining the permission of the research subjects, the researcher entered the classroom in advance according to the teaching period and the teaching site provided by the research subjects, chose a location in the middle and back of the classroom that was conducive to observation, covered the complete teaching area, had good sound pickup, did not interfere with the classroom teaching process, and was relatively secluded, and filled in the basic information on the observation record sheet of the current classroom in order to make preparations for the classroom observation. Since this study focuses on the knowledge visualization teaching practices of university English teachers, the various forms of visual representations that teachers use in the classroom to carry foreign language knowledge are included in the scope of this study, as well as the information technology tools, multimedia language learning platforms, and related resources used by the teachers.

4.2.2. Supporting information collection. The auxiliary information data in this study include two categories: textual information data and audiovisual material data. Textual material data mainly includes the paper or electronic documents related to teaching of the research participants, such as course textbooks, teachers' books, some teaching reference materials, course teaching plans, course lesson plans, teaching courseware, and so on. It also includes the teaching logs, personal teaching reflections, and some online review text data of the research participants. The data of audiovisual materials mainly include a series of multimedia language teaching resources such as videos, audios, photos, drawing images, and so on, which are closely related to teaching and learning, and are used by the research participants in the preparation of lessons, the distribution and collection of students' homework, and the process of teaching in the classroom. The above data will be provided by the research participants on the condition that the researcher informs them of the confidentiality and anonymity measures. Or, with the permission of the participants, the researcher redrew the data based on the original data recorded by the video recording equipment.

After organizing, classifying, extracting, and reproducing the above data, the researcher will provide feedback to the research subjects with the edited and summarized data, and after obtaining their approval and consent, the researcher will unify and code the above data with the data from the supporting materials.

4.3. Principal component analysis process

The raw data were first standardized to eliminate the effect of different indicator scales. The standardization formula is:

$$x^*_{ij} = \frac{x_{ij} - \overline{x}_j}{s_j} \tag{9}$$

Where x^*_{ij} is the standardized variable value, x_{ij} is the original variable value, $\overline{x}_j$ is the mean of the j th variable, and s_j is the standard deviation of the j th variable. Then the covariance matrix or correlation matrix of the standardized data was calculated, followed by solving its eigenvalues and

eigenvectors. The number of principal components was determined according to the size of the eigenvalues, and the first m principal components with a variance contribution rate of 85% or more were selected. Determine the principal component loading matrix, calculate the score of each student on each principal component, and finally construct a comprehensive evaluation function based on the principal component scores and the variance contribution rate:

$$F = \sum_{j=1}^{m} \eta_j z_j \tag{10}$$

Where F is the comprehensive score of students' language proficiency, η_j is the variance contribution rate of the j rd principal component, and z_j is the score of the j th principal component. With this comprehensive evaluation function, the effect of students' language proficiency enhancement in task-based teaching of English in IT-assisted colleges and universities can be quantitatively measured.

5. Empirical studies

5.1. Subjects and experimental design

In order to fully reflect the role of information technology-assisted task-based English teaching method on the improvement of students' language ability, this paper enlarges the sample capacity when selecting the research subjects, and selects a total of 480 students in 10 classes majoring in English in a university, which are randomly divided into six groups of 80 students each. Among them, the first two groups serve as the experimental group, noted as E1 and E2, adopting this paper's task-based English teaching model based on SPOC information technology. The remaining four groups serve as the control group, noted as C1, C2, C3 and C4, adopting the traditional English teaching mode. Before the experiment, the students in the six groups were given a comprehensive test of language ability, including vocabulary, grammar, listening, speaking, reading and writing. Figure 2 shows the statistics of the comprehensive language ability of the 6 groups of students obtained after the test according to the comprehensive evaluation function constructed in this paper. As can be seen from the figure, before the beginning of the experiment, the mean values of the comprehensive scores of language proficiency of the 6 groups of students are 56.06, 53.29, 54.08, 55.85, 52.07 and 55.17, respectively, and there is no significant difference.

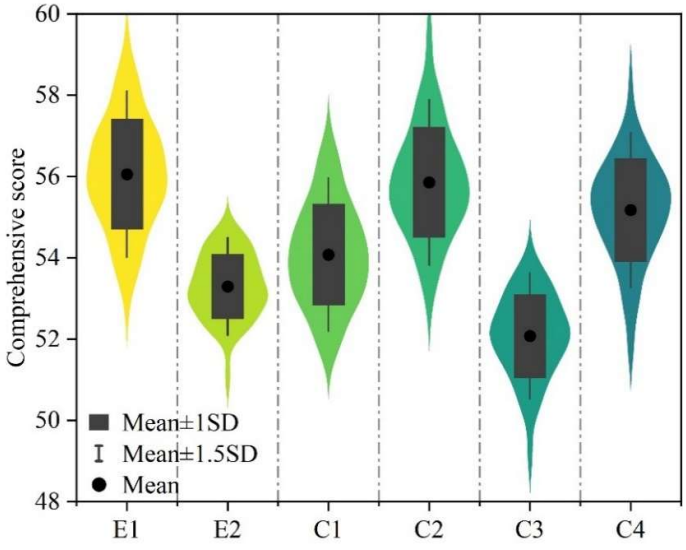

Fig. 2. Comprehensive language ability score statistics

The experimental cycle is one semester, and the experimental group uses multimedia teaching resources, MOOC online learning platform, task management system and other information technology means to carry out task-based English teaching, while the control group adopts the traditional classroom lectures, post-class homework exercises and other ways of teaching.

5.2. Data collection and analysis

At the end of the experiment, data were collected according to the measurement index system constructed above. For the students in the experimental group and the control group, their language knowledge test scores, language skills assessment scores, learning attitudes and strategies related data and task completion quality evaluation data were collected respectively, and all of them were assigned scores according to the percentage system. After the collected data were sorted out, principal component analysis was used to construct the measurement model of language proficiency enhancement for the experimental group and the control group students respectively, and the composite score of each student was calculated.

5.3. Results and analysis

5.3.1. Primary data collection. According to the method of data collection and organization, the raw data of the indicators of the effect of language proficiency improvement of the students in the experimental group and the control group after the experiment are obtained, and the results are shown in Table 2. From the table, it can be seen initially that after the experiment, the language proficiency test scores of the students in the experimental group are significantly improved compared with those of the students in the control group. In order to more profoundly and intuitively test the effect of IT-assisted task-based English teaching on students' language proficiency enhancement, the later paper utilizes the measurement model based on principal component analysis constructed in this paper to quantify the students' language proficiency enhancement.

Table 2. Raw data on student language ability

Indicator	E1	E2	C1	C2	C3	C4
X1	99.27	94.52	65.78	75.02	64.82	76.46
X2	92.14	86.99	67.00	64.71	65.20	75.09
X3	80.76	90.16	69.47	73.41	71.82	70.86
X4	99.35	87.00	62.04	71.57	60.51	76.99
X5	94.88	82.25	60.12	71.61	70.92	67.54
X6	85.54	91.51	70.66	79.57	78.65	75.47
X7	82.91	92.98	79.34	64.67	71.41	74.70
X8	93.76	81.72	73.00	70.83	63.68	79.75
X9	92.70	84.02	60.48	62.63	62.65	65.72
X10	97.94	89.71	70.8	65.17	63.45	77.73
X11	87.81	81.02	66.87	74.40	73.31	65.83
X12	86.74	90.34	75.06	66.58	69.31	70.21
X13	86.58	96.38	69.11	79.25	71.11	71.48

Table 3 shows the results obtained after normalization of the original data. Data normalization can effectively avoid the large error between the results and the actual situation caused by the different scale and other reasons, so as to improve the accuracy of this paper's model in measuring students' language proficiency improvement.

Table 3. Normalized data

Indicator	E1	E2	C1	C2	C3	C4
X1	0.084	0.082	0.074	0.082	0.073	0.081
X2	0.078	0.076	0.075	0.070	0.074	0.079
X3	0.068	0.078	0.078	0.080	0.081	0.075
X4	0.084	0.076	0.070	0.078	0.066	0.081
X5	0.083	0.071	0.068	0.078	0.080	0.071
X6	0.072	0.080	0.079	0.087	0.089	0.081
X7	0.070	0.081	0.089	0.070	0.081	0.079
X8	0.079	0.071	0.082	0.077	0.072	0.084
X9	0.079	0.073	0.068	0.068	0.071	0.069
X10	0.083	0.078	0.080	0.071	0.072	0.082
X11	0.074	0.071	0.075	0.081	0.083	0.069
X12	0.073	0.079	0.084	0.072	0.078	0.074
X13	0.073	0.084	0.078	0.086	0.080	0.075

5.3.2. Principal component analysis. Table 4 shows the initial solution for the principal component analysis, showing the commonality data for all quantities. Column 2 of the table shows the initial commonality. The initial common are obtained after considering all the components and factors affecting the students' language proficiency improvement to get the variance estimates for each variable. In the correlation analysis, the commonality value obtained by extracting the factors using the system's default principal component method is constant at 1, which means that the variance of all the indicator variables after standardization is 1, and all the variance of the original indicator variables can be explained. The purpose of principal component analysis is to find the number of factors smaller than the number of variables, so it is not possible to take all the eigenvalues. Column 3 in the table shows the common degree of the extracted eigenvalues, which is the common degree under the condition that the specified eigenvalues are greater than. It can be seen that the common degree of the 13 variables is basically above 90%, i.e., there is very little information loss in these indicator variables, and the indicator variables can all be explained by the principal components. Therefore, the effect of principal component extraction is better. If there exists a low value of common degree among them under the principal component extraction method, you can re-select other extraction methods or re-specify the criteria for extracting eigenvalues.

Table 4. Principal component analysis common

Indicator variable	Initial common	Extraction common
X1	1.000	0.923
X2	1.000	0.912
X3	1.000	0.990
X4	1.000	0.992
X5	1.000	0.905
X6	1.000	0.928
X7	1.000	0.973
X8	1.000	0.930
X9	1.000	0.903
X10	1.000	0.975
X11	1.000	0.917
X12	1.000	0.950
X13	1.000	0.933

From the above common degree of principal component analysis, it can be seen that the correlation coefficients between the 13 related indicators of the effect of students' language proficiency enhancement are all large, thus indicating high correlation, and in order to avoid overlapping

information and simplify the model, the method of principal component analysis can be used to extract the principal component common factor from the many factors, and then carry out the correlation study. Using statistical software to organize the collected data, the results of principal component analysis are shown in Table 5, where T, V, and C represent the total, the percentage of variance, and the cumulative contribution rate, respectively. From the table, it can be seen that the eigenvalue of the first principal component is 29.873, which explains 80.361% of the total variance of the 13 original variables, and the cumulative variance contribution rate is 80.361%. The eigenvalue of the second principal component is 2.854, explaining 8.482% of the total variance of the 13 original variables with a cumulative variance contribution of 88.843%. The eigenvalue of the third principal component is 1.849, explaining 5.346% of the total variance of the 13 original variables with a cumulative variance contribution of 94.189%. The fourth principal component has an eigenvalue of 1.182 and explains 2.870% of the total variance of the 13 original variables with a cumulative variance contribution of 97.059%. All 13 factors were presented in the initial solution and the original variables were explained 100% from the 8th onwards.

Table 5. The principal component analysis reveals the total variance

Componen t	Initial eigenvalues			Extract sum of squares			Rotational squares		
	T	V/%	C/%	T	V/%	C/%	T	V/%	C/%
1	29.873	81.473	81.473	29.873	81.473	81.473	17.834	46.837	46.837
2	2.854	7.784	89.257	2.854	7.784	89.257	13.291	40.296	87.133
3	1.849	5.043	94.300	1.849	5.043	94.300	2.027	6.943	94.076
4	1.182	3.224	97.524	1.182	3.224	97.524	1.398	3.178	97.524
5	0.373	1.017	98.541						
6	0.295	0.805	99.346						
7	0.162	0.442	99.788						
8	0.078	0.212	100						
9	0.000	0	100						
10	0.000	0	100						
11	0.000	0	100						
12	0.000	0	100						
13	0.000	0	100						

Figure 3 Fragmentation diagram for principal component analysis. The use of the rubble diagram of principal component analysis can help to determine the optimal number of factors for the effectiveness of students' language proficiency improvement. In the gravel plot, the horizontal coordinates indicate the number of principal components and the vertical coordinates indicate the eigenvalues. From the graph, it can be seen that the eigenvalues of the first 4 principal components are large, and from the 5th the eigenvalues become very small, and the factor eigenvalues become flat, that is to say, the first 4 factors contribute a lot to the explanatory variables, so it is most appropriate to extract the 4 principal components in the factor analysis. They are first recorded as male factor F1, male factor F2, male factor F3 and male factor F4 respectively.

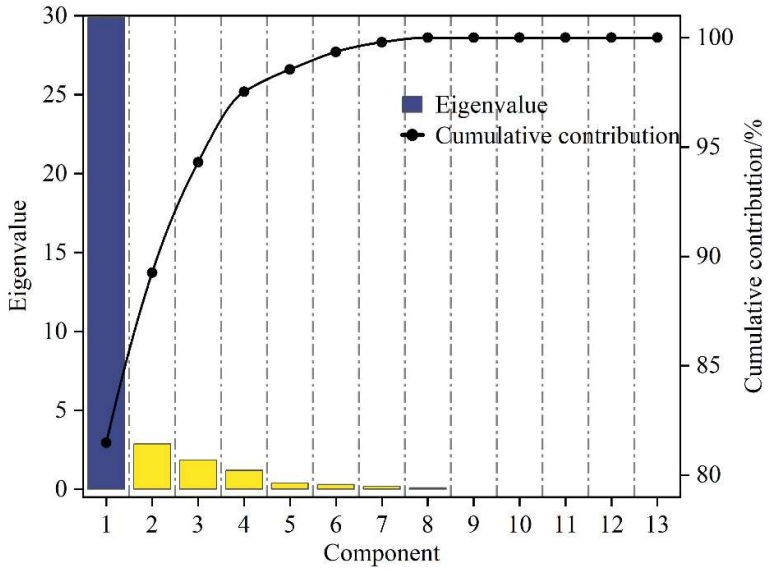

Fig. 3. The rubble of principal component analysis

Figure 4 shows the factor loading matrix before the rotation of the effect of the students' language proficiency improvement, and through the factor loading coefficients of the individual principal components in the matrix, it is possible to represent the initial individual factors as linear combinations of the common factors, such as:

$$X1 = 0.829F1 + 0.039F2 + 9.927F3 + 0.006F4$$

...

$$X13 = 0.894F1 + 0.022F2 - 0.004F3 + 0.01F4$$

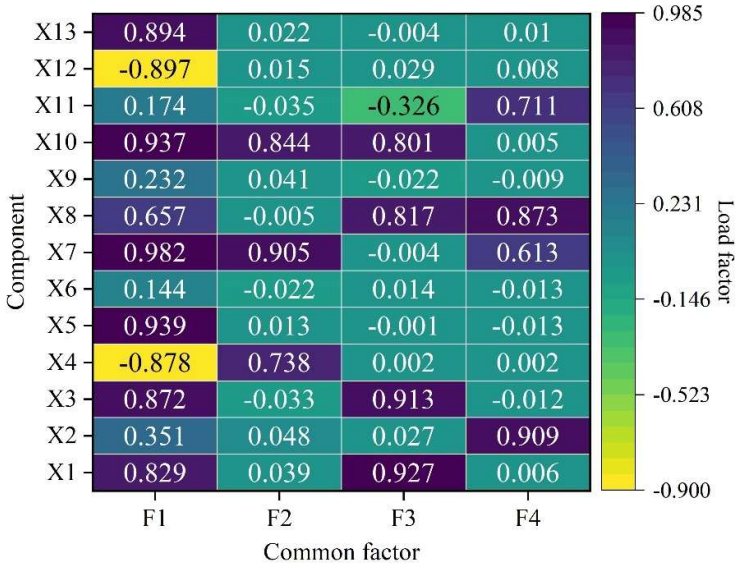

Fig. 4. Factor load matrix

Figure 5 shows the rotated principal component factor loading coefficients, using the rotated loading coefficients, the meaning of the principal components can be better low revealed. From the factor loading matrix in Figure 4, it can be seen that each factor has a certain loading on each original variable, and it is difficult to see the pattern, thus in order to better study the extracted public factors and give each public factor a practical meaning, the factor loading array is rotated using the method of rotating the maximum variance. It can be seen that the rotated individual male factors have larger

loadings on the original factors associated with them respectively, and smaller loadings on the rest of the factors, therefore, individual male factors can be named according to their loadings. From the figure, it can be seen that the male factor F1 has larger loadings on the original factors X1 and X2, and smaller loadings on the rest of the factors, and the factors with larger loadings can all be categorized as factors in terms of linguistic knowledge, and thus the male factor F1 can be named linguistic knowledge. The male factor F2 has larger loadings on the original factors X3, X4, X5 and X6, smaller loadings on the rest of the factors, and the factors with larger loadings can all be categorized as factors in language skills, thus the male factor F2 can be named as language skills. According to this naming method, F3 and F4 can be named Learning Attitude and Strategies and Quality of Task Completion, respectively.

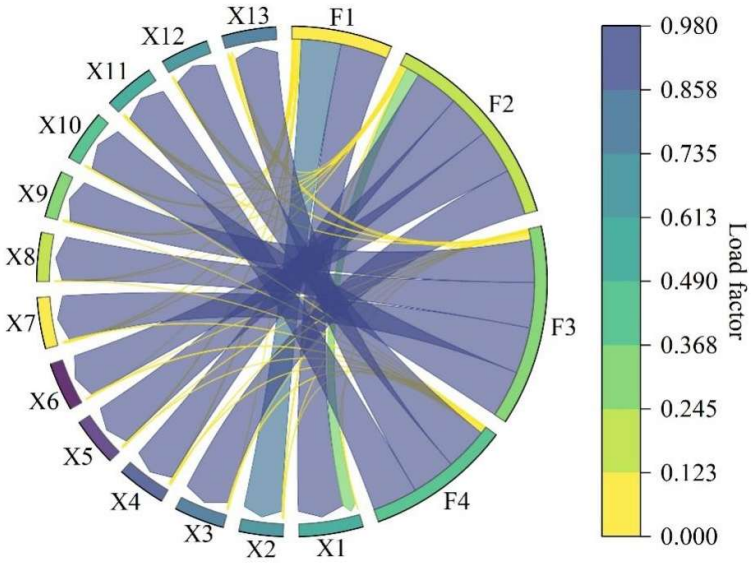

Fig. 5. The factor load coefficient of the rotation

From the above factor loading coefficients can be obtained from the loadings of each male factor on each of the original initial factors, the main factors are analyzed, and then the factor score coefficient matrix can be obtained as shown in Table 6. According to the matrix of factor score coefficients can be obtained from the relationship between each male factor and each of the original factors, each male factor is expressed as a linear combination of the original factors in the form of the following:

$$F1 = 0.056X1 + 0.078X2 + 0.012X3 + 0.011X4 + 0.004X5 - 0.021X6 + 0.006X7$$
$$+0.018X8 + 0.003X9 - 0.001X10 + 0.011X11 + 0.009X12 - 0.013X13,$$
$$F2 = 0.013X1 + 0.017X2 + 0.059X3 + 0.062X4 + 0.055X5 + 0.079X6 + 0.003X7$$
$$-0.002X8 - 0.005X9 + 0.008X10 + 0.007X11 - 0.013X12 + 0.002X13,$$
$$F3 = -0.004X1 + 0.002X2 - 0.006X3 + 0.002X4 + 0.012X5 + 0.008X6 + 0.052X7$$
$$+0.067X8 + 0.046X9 + 0.073X10 - 0.004X11 + 0.003X12 - 0.008X13$$
$$F4 = 0.008X1 + 0.005X2 + 0.007X3 - 0.004X4 + 0.001X5 + 0.009X6 - 0.002X7$$
$$+0.001X8 + 0.007X9 = 0.008X10 + 0.058X11 + 0.064X12 + 0.049X13$$

Table 6. Score coefficient matrix of the principal component

Indicator variable	Principal component			
	F1	F2	F3	F4
X1	0.056	0.013	-0.004	0.008
X2	0.078	0.017	0.002	0.005
X3	0.012	0.059	-0.006	0.007
X4	0.011	0.062	0.002	-0.004
X5	0.004	0.055	0.012	0.001
X6	-0.021	0.079	0.008	0.009
X7	0.006	0.003	0.052	-0.002
X8	0.018	-0.002	0.067	0.001
X9	0.003	-0.005	0.046	0.007
X10	-0.001	0.008	0.073	-0.008
X11	0.011	0.007	-0.004	0.058
X12	0.009	-0.013	0.003	0.064
X13	-0.013	0.002	-0.008	0.049

According to the above analysis combined with the comprehensive evaluation function of students' language proficiency constructed in this paper, the comprehensive scores of students' language proficiency were finally calculated as shown in Table 7, and the comprehensive scores were processed in percentage system for the sake of easy comparison to reflect the status of students' language proficiency improvement. The data in the table shows that the students in the two experimental groups E1 and E2 have the same composite score of language proficiency, which is the highest score of 86.31, while the rest of the four control groups, C1, C2, C3, and C4, have the same composite score of language proficiency, which is 58.98, 59.93, 61.91, and 60.26, respectively. From the results, it can be seen that after the experiment, there was a significant difference in the comprehensive scores of language proficiency between the students in the experimental group and the students in the control group. The comprehensive scores of the students in the experimental group increased by 46.34%, 44.02%, 39.41% and 43.23% respectively compared with the control group, and 53.96% and 61.96% respectively compared with the pre-experiment. It shows the effectiveness of the task-based English teaching model constructed on the basis of SPOC technology and the significant enhancement of students' language proficiency by information technology-assisted task-based English teaching in colleges and universities.

Table 7. Student language ability comprehensive score

Group	F1	F2	F3	F4	Comprehensive score
E1	0.892	0.868	0.836	0.974	86.31
E2	0.896	0.836	0.921	0.803	86.31
C1	0.604	0.632	0.611	0.629	58.98
C2	0.617	0.608	0.637	0.619	59.93
C3	0.636	0.664	0.643	0.602	61.91
C4	0.616	0.627	0.659	0.668	60.26

6. Conclusion

This paper utilizes information technology to assist task-based teaching and introduces it into English teaching in colleges and universities. It adopts principal component analysis to process the data of students' language proficiency related indexes, and constructs the measurement model of students'

language proficiency improvement to calculate the comprehensive scores of the experimental group and the control group after the experiment. The comprehensive scores of language proficiency of students in the experimental group increased by 46.34%, 44.02%, 39.41% and 43.23% respectively compared with those of the control group, and there was a significant difference between the language proficiency of students in the two groups after the experiment. Meanwhile, the mean language proficiency scores of the students in the experimental group before the experiment were 56.06 and 53.29 respectively, while the scores after the experiment were 86.31, with an increase of 53.96% and 61.96% respectively. It intuitively shows that the information technology combined with the task-based English teaching mode has an important role in the improvement of students' language proficiency.

Funding

Funded by Project of Beijing Institute of Petrochemical Technology: 2024 Graduate Education and Teaching Reform Research and Practice---Curriculum Construction and Teaching Reform (YJ24-103) and funded by the Innovation and Entrepreneurship Training Program for College Students of 2025.

References

[1] Jackson, D. O. (2022). Task-Based Language Teaching. Elements in Language Teaching.

[2] Halici Page, M., & Mede, E. (2018). Comparing task-based instruction and traditional instruction on task engagement and vocabulary development in secondary language education. Journal of Educational Research, 111(3)

[3] Ellis, R. (2017). Position paper: Moving task-based language teaching forward. Language teaching: The international abstracting journal for language teachers and applied linguistics, 50(4), 507-526.

[4] Bhandari, L. P. (2020). Task-Based Language Teaching: A Current EFL Approach. Advances in Language and Literary Studies, 11(1), 1.

[5] Sholeh, M. B. (2020). Implementation of task-based learning in teaching English in Indonesia: Benefits and problems. Language Circle: Journal of Language and Literature, 15(1), 1-9.

[6] Ji, Y., & Pham, T. (2018). Implementing task-based language teaching (tblt) to teach grammar in english classes in china: using design-based research to explore challenges and strategies. Innovation in Language Learning and Teaching, (1), 1-14.

[7] Adiantika, H. N., & Purnomo, H. (2018). The implementation of task-based instruction in efl teaching speaking skill. Indonesian EFL Journal, 4(2), 12.

[8] Nychkalo, N., Jinba, W., Lukianova, L., Paziura, N., & Muranova, N. (2020). Use of Task-Based Approach in Teaching Vocabulary to Business English Learners at University. Advanced Education, 16, 98-103.

[9] Liu, Y., & Ren, W. (2021). Task-based language teaching in a local EFL context: Chinese university teachers' beliefs and practices. Language Teaching Research, 1, 17.

[10] Sholeh, M. B., Salija, K., & Nur, S. (2021). TASK-BASED LEARNING IN ENGLISH AS A FOREIGN LANGUAGE (EFL) CLASSROOM: WHAT, HOW AND WHY?. Getsempena English Education Journal, 8(1), 134-146.

[11] Roy, D. (2017). Task-based EFL language teaching with procedural information design in a technical writing context. Cogent Education, 4(1).

[12] Nobre, A. M. D. J. F. (2018). Multimedia Technologies and Online Task-Based Foreign Language Teaching-Learning. Tuning Journal for Higher Education, 5(2), 75-97.

[13] Rodríguez-Peñarroja, M. (2022). INTEGRATING PROJECT-BASED LEARNING, TASKBASED LANGUAGE

TEACHING APPROACH AND YOUTUBE IN THE ESP CLASS: A STUDY ON STUDENTS'MOTIVATION. Teaching English with Technology, 22(1).

[14] Masuram, J., & Sripada, P. N. (2020). Developing spoken fluency through task-based teaching. Procedia Computer Science, 172, 623-630.

[15] Lin, C. C. (2021, July). The impact of technology-assisted task-based language teaching on self-regulated learning. In 2021 International Conference on Advanced Learning Technologies (ICALT) (pp. 285-287). IEEE.

[16] Tiu, J., Groenewald, E., Kilag, O. K., Balicoco, R., Wenceslao, S., & Asentado, D. (2023). Enhancing Oral Proficiency: Effective Strategies for Teaching Speaking Skills in Communication Classrooms. Excellencia: International Multi-disciplinary Journal of Education (2994-9521), 1(6), 343-354.

[17] Chen, K. T. C. (2021). The Effects of Technology-Mediated TBLT on Enhancing the Speaking Abilities of University Students in a Collaborative EFL Learning Environment. Applied Linguistics Review, 12(2).

[18] Tseng, M. F., & Geng, Z. (2021). Implementing virtual reality–enhanced tasks in Chinese language teaching. In Contextual Language Learning: Real Language Learning on the Continuum from Virtuality to Reality (pp. 91-118). Singapore: Springer Singapore.

[19] Cardot, H., & Degras, D. (2018). Online Principal Component Analysis in High Dimension: Which Algorithm to Choose?. HAL, 2018.

[20] Yildirim, I. (2017). The effects of gamification-based teaching practices on student achievement and students' attitudes toward lessons. The Internet and Higher Education, 100(33), 86-92.

[21] Qiao Zhou. (2024). Research on Computer-Assisted Task-Based Instruction in College English Writing Teaching. Journal of Educational Research and Policies(11),1-4.

[22] Liu Lei. (2024). Empirical Research on SPOC Teaching in College Sports Course Reform: A Case Study of Basketball Courses. Curriculum and Teaching Methodology(8).

[23] Zhiqiang Liu. (2024). Research on the evaluation system and teaching optimization effect of animation aesthetics education based on principal component analysis method. Applied Mathematics and Nonlinear Sciences(1).