ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 2 - 2026

Research on the Adaptation Mechanism of Dynamic Assessment and Deep Learning Objectives and the Optimization of Teaching Efficiency in English Teaching

Lili Niu[1,*] *and Wenchao Fan*[2]

[1] School of Foreign Language, Xinxiang Institute of Engineering, Xinxiang 453700, Henan, China

[2] Department of Computer Science and Engineering, Cangzhou Normal University Cangzhou 061001, China

SUMMARY: *In order to improve the accuracy and feedback efficiency of dynamic assessment results of English teaching, a teaching effectiveness optimization model for deep learning target adaptation was proposed. The model collects data such as platform logs, writing texts, oral audio, reading tests and interactive feedback, and completes standardization processing, feature extraction and learning ability portrait. Then, the index system of language foundation, discourse understanding, expression generation, thinking processing, interaction and collaboration and self-regulation are constructed, and the mapping relationship between dynamic evaluation index and deep understanding, comprehensive expression, transfer application and other goals is established. The experiment was carried out with 96 students for 16 weeks. The comprehensive score of the experimental class increased from 71.36 to 84.21, the feedback acceptance rate reached 84.36%, the evaluation time was reduced to 8.4 minutes, and the average error was 3.18%. The results show that the proposed model can improve the efficiency of English teaching evaluation and the achievement of deep learning goals.*

KEYWORDS: *dynamic evaluation; Deep learning objective; English teaching; Optimization of Teaching Efficiency*

1 Introduction

In recent years, English teaching evaluation has gradually shifted from summative achievement judgment to learning process diagnosis and ability development tracking. Traditional tests can reflect students' vocabulary, grammar, reading and writing performance at a specific time point, but it is difficult to reveal learners' cognitive regulation, strategy use and language transfer ability in the process of task completion. Dynamic assessment emphasizes the continuous relationship of "evaluation-feedback-re-performance", which can observe students' potential development level in teaching intervention and provide a more detailed basis for improving English teaching efficiency. Kushki et al. introduced dynamic assessment into EFL argumentative writing teaching and proved that mediated feedback could help reveal learners' real difficulties in writing development [1]. Kushki et al. further analyzed the process of dynamic writing assessment from the perspective of task response, and pointed out that the assessment should not stop at the score of text results, but should focus on the

*lunwen20232023@163.com

https://doi.org/10.65102/is2026981 Published: 30 April, 2026

change of students' ability in prompting, revision and re-expression [2].

With the introduction of mobile learning platforms, online writing systems, speech recognition tools and classroom interactive platforms into English teaching, students 'learning behavior trajectories, text output, speech performance, feedback response and task completion time have gradually formed computable data. Rassaei carried out dynamic evaluation of request expression teaching through mobile media, indicating that digital environment can enhance the immediacy and individual adaptability of feedback [3]. Kafipour and Khoshnood explored the role of dynamic assessment feedback on the development of oral English ability from the perspective of differences in learning styles, and showed that there were obvious differences in the acceptance of feedback types among different learners [4]. Zadkhast et al. compared the effect of concurrent and cumulative group dynamic assessment in the development of reading micro-skills, suggesting that English teaching evaluation should take into account both individual performance and group learning process [5]. Randall and Urbanski developed a computerized dynamic assessment program for second language grammar, indicating that automated assessment tools can assume the functions of diagnosis, prompt and record in grammar teaching [6].

Existing studies provide rich experience for English dynamic assessment, but there are still shortcomings in deep learning target adaptation. The deep learning objective here does not simply refer to neural network training, but to higher-order learning objectives such as understanding construction, meaning transfer, critical expression, collaborative communication and self-reflection in English learning. Some studies focus on grammar mastery, writing revision or reading micro-skill improvement, but few of them correspond the evaluation indicators to the deep language ability structure [7-9]. At the same time, although teachers' reflection, learners' autonomy and classroom communication ability have been included in the perspective of dynamic evaluation, there is still a lack of stable mechanism for how the evaluation results are transformed into the optimization of teaching strategies [10]. Suri pointed out that the cultivation of deep learning ability in English classroom requires teachers to form a continuous support among task design, language application and thinking training [11], which provides a direction for the reconstruction of dynamic assessment indicators.

The development of artificial intelligence technology provides a new computational foundation for English teaching effectiveness evaluation. Related studies have shown that artificial intelligence can be used in English learning resource recommendation, automatic writing evaluation, speech recognition, learning behavior analysis and personalized feedback generation [12-15]. However, if AI tools are only used as raters or error correctors, it is easy to cause fragmentation of evaluation results, and it is difficult to serve the continuous achievement of deep learning goals. Based on this, this paper focuses on the adaptation relationship between dynamic assessment and deep learning objectives in English teaching, and constructs a model of learning process data collection, learning performance analysis, dynamic assessment index classification, goal mapping and teaching effectiveness optimization. With the help of natural language processing, learning analysis and intelligent feedback mechanism, this study attempts to link students 'language performance and behavior data with teachers 'teaching adjustment, and provide method support for English teaching evaluation from result judgment to process diagnosis, ability promotion and effectiveness optimization.

2 The application of dynamic assessment in deep learning target adaptation in English teaching

The application of dynamic assessment in English teaching does not simply convert students 'stage results into grades, but integrates the learning process, teacher's feedback, students 'revision and ability re-performance into the same evaluation chain. The goal of English deep learning emphasizes the comprehensive development of students in language understanding, meaning construction, cross-context expression, critical thinking and autonomous regulation. Therefore, the evaluation system needs to break through the static judgment method of a single test and shift to continuous perception, dynamic diagnosis and immediate intervention. The development of artificial intelligence and learning analysis technology enables text assignments, voice exercises, reading tracks, online tests, classroom interactions and feedback responses in English classrooms to be transformed into computable data, providing technical conditions for the establishment of an adaptation relationship between dynamic assessment and deep learning objectives [16, 17].

In the English teaching scenario, the core task of dynamic assessment is to identify the gap between a student's current performance and the target ability, and generate the corresponding teaching support according to the gap type. If a student has a high reading accuracy rate but loses significant points in reasoning questions, the system needs to determine whether the problem comes from vocabulary comprehension, discourse logic, insufficient background knowledge or improper strategy use. If the students' writing grammar errors are reduced but the argumentation level is weak, the evaluation results should point to the deep ability of idea development, evidence organization and discourse cohesion. Based on this idea, this paper divides English learning data into four categories: behavior data, language output data, interactive feedback data and stage evaluation data, and forms a dynamic evaluation application framework for deep learning goals through data cleaning, feature extraction, target mapping and feedback update, as shown in Figure 1.

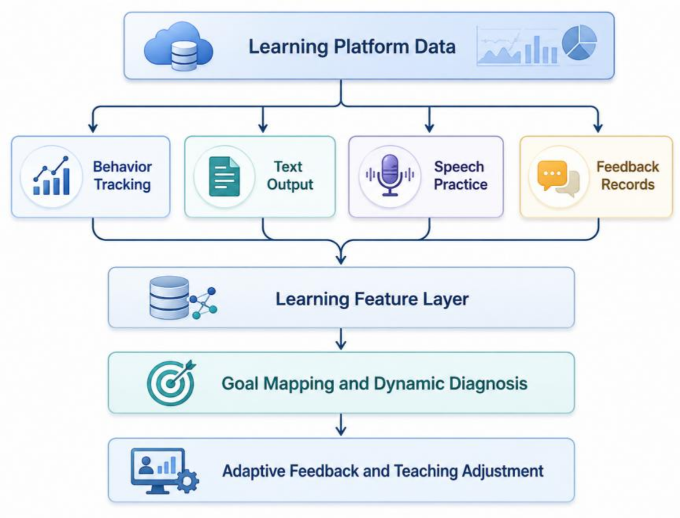

Figure 1: Dynamic assessment application framework for English deep learning objectives

2.1 English learning process data collection and learning performance analysis

Data collection in English learning process is the basis of dynamic assessment mechanism. In this paper, the learning process is understood as a continuous collection of student behaviors in pre-class preview, classroom participation, after-class training, feedback revision and stage assessment. The data sources mainly included online learning platform logs, English writing submission records, oral following and free expression audio, reading comprehension answer tracks, classroom interaction records and teacher feedback texts. The data from different sources have structural differences, and the platform log is mostly manifested as time, frequency and path. The writing and speaking data belong to unstructured language output, and the test data has clear items and scores. To ensure the stability of subsequent analysis, it is necessary to encode the multi-source data uniformly. Let the learning process data of the ith student in the TTH learning cycle be expressed as follows.

$$X_i^t = \{B_i^t, W_i^t, S_i^t, R_i^t, F_i^t\} \tag{1}$$

Among them, B_i^t represents learning behavior data, W_i^t represents writing text data, S_i^t represents oral performance data, R_i^t represents reading and testing data, and F_i^t represents teacher or system feedback data. This expression can be used to analyze both the explicit achievement and implicit process in English learning, and avoid judging students' ability level only by one test score.

After the data collection is completed, the raw data needs to be cleaned and standardized. For the missing learning records, the system marked them according to the type of learning task instead of deleting them directly. For abnormal login time, repeated submission and invalid speech segments, it is processed by threshold detection and manual sampling review. Due to the different dimensions of different indicators, this paper uses the range standardization method to map the data to a unified interval:

$$z_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j) + \varepsilon} \tag{2}$$

where x_{ij} is the original value of the ith student on the JTH index, $\max(x_j)$ and $\min(x_j)$ represent the maximum and minimum value of this index, respectively, and ε is a smoothing term to prevent the denominator from being zero. Through this processing, indicators such as learning frequency, task completion rate, writing score, oral fluency and feedback adoption rate can be entered into the same analysis space.

English learning performance analysis needs to take into account both language accuracy and learning development. This paper establishes the learning performance vector from the five dimensions of knowledge acquisition, language expression, thinking processing, interactive participation and self-regulation:

$$P_i^t = (K_i^t, E_i^t, C_i^t, I_i^t, A_i^t) \tag{3}$$

K_i^t represents the mastery of vocabulary, grammar, and syntactic structure; E_i^t represents the quality of writing and speaking; C_i^t represents the ability of understanding reasoning, opinion organization and cross-context transfer. I_i^t represents classroom interaction, peer collaboration and feedback response; A_i^t represents learning plan, revision behavior, and autonomic regulation ability. This vector enables dynamic assessment to shift from individual skill analysis to comprehensive ability profiling.

In the specific calculation, the system combines automatic writing evaluation, speech recognition, reading behavior analysis and feedback text annotation technology to evaluate the students 'performance in multiple dimensions. For example, writing performance not only focuses on the number of spelling and grammatical errors, but also analyzes topic relevance, paragraph cohesion, argumentation adequacy, and revision magnitude. Oral performance not only focuses on pronunciation accuracy, but also considers the stability of speaking rate, pause distribution and expression integrity. In addition to the correct rate, the reading performance also recorded the number of questions looking back, the dwell time and the type of reasoning questions error. The resulting comprehensive value of learning performance can be expressed as follows.

$$L_i^t = \sum_{j=1}^{m} \omega_j z_{ij} \tag{4}$$

Here, L_i^t is the comprehensive learning performance value of the ith student in the TTH cycle, ω_j represents the JTH index weight, and m is the number of indicators involved in the calculation. The weight setting was determined by teaching objectives, task types and expert evaluation, so as to avoid the system over-relying on easy to quantify indicators and ignoring the deep learning objectives.

Dynamic evaluation also needs to embody "change after feedback". Therefore, this paper introduces the learning gain metric, which is used to analyze the change in student performance after receiving prompts, examples, explanations or peer reviews:

$$G_i^t = \frac{L_i^{t+1} - L_i^t}{1 - L_i^t + \varepsilon} \tag{5}$$

Here, G_i^t represents the relative improvement of students in adjacent learning cycles. If the original performance of students is low but the improvement is obvious after feedback, it means that their potential development space is large. If students have high scores but insufficient gains, it is necessary to further observe whether there are problems of insufficient task challenges or solidified learning strategies. The index is able to help teachers distinguish between the two different states of "current low performance" and "arrested development", making the teaching intervention more accurate.

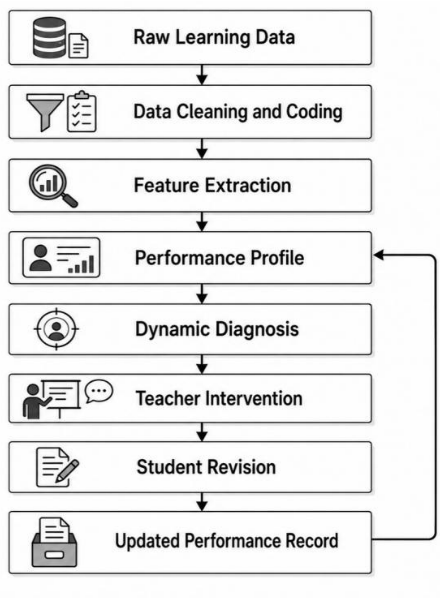

Figure 2: Learning performance analysis and feedback update process

Figure 2 shows the learning performance analysis and feedback update process. The system does not take the rating as the end point, but continues to pass the diagnosis results to the teacher intervention and student revision sessions. Teachers can design hierarchical tasks for different students according to the performance portraits output by the system. For example, for students with weak vocabulary foundation but positive interaction, we should strengthen chunk input and sentence pattern support. For students with less grammatical errors but insufficient ideas, we should increase the training of discussion, example organization and critical expression. For students with high task completion rate but insufficient revision depth, writing version comparison and reflection journal were introduced. In this way, the dynamic assessment results can be translated into specific teaching actions, rather than staying at the data presentation level.

2.2 English Teaching dynamic assessment index classification and deep learning target mapping

Because the process of English teaching has the characteristics of multiple types of tasks, large differences in learning behaviors, complex forms of language output and continuous changes in feedback results, dynamic evaluation indicators cannot be simply divided according to the types of test questions. If only listening, speaking, reading and writing are scored separately, the evaluation results are easy to stay at the skill surface, and it is difficult to reflect students' actual development status in deep learning goals such as understanding construction, meaning transfer, critical expression and autonomous regulation. Therefore, after completing the data collection and performance analysis of the English learning process, it is necessary to classify the dynamic evaluation indicators, and further establish the mapping relationship between the indicator categories and the deep learning objectives. The role of this process is to transform scattered learning data into interpretable ability evidence, so that teachers can judge the source of students 'learning difficulties based on the evaluation results, and adjust the teaching tasks, feedback methods and support intensity accordingly.

2.2.1 Classification of dynamic assessment indicators for English teaching

The classification of dynamic assessment indicators of English teaching is not simply data grouping, but according to the occurrence logic of English learning activities, the data from different sources are grouped into indicators that can explain the development of learning ability. In this paper, the dynamic assessment indicators are divided into five categories: language basic indicators, task performance indicators, cognitive processing indicators, interactive feedback indicators and learning regulation indicators. The basic language indicators mainly include vocabulary mastery, grammar accuracy, syntactic complexity and speech clarity. The task performance indicators included reading comprehension accuracy, writing completion quality, oral expression integrity and listening information extraction effect. The cognitive processing index mainly reflected students' performance in reasoning judgment, opinion organization, discourse cohesion and cross-context transfer. The interactive feedback index is used to describe students' responses to teacher prompts, system suggestions and peer evaluations. The learning regulation index focused on task planning, revision frequency, error review, reflection recording and learning rhythm stability. Let the set of dynamic evaluation indicators of the ith student in the TTH evaluation cycle be as follows.

$$D_i^t = \{d_{i1}^t, d_{i2}^t, \cdots, d_{im}^t\} \tag{6}$$

Here, m represents the number of indicators participating in the dynamic evaluation, and

d_{ij}^{t} represents the standardized result of the JTH indicator of student i in period t. In order to reflect the difference in the contribution of different indicators to the ability diagnosis, this paper sets up the index classification weight matrix:

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1m} \\ c_{21} & c_{22} & \cdots & c_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ c_{q1} & c_{q2} & \cdots & c_{qm} \end{bmatrix} \tag{7}$$

where q represents the number of dynamic evaluation index categories, and c_{rj} represents the weight of the JTH index belonging to the RTH index. If an index contributes more to a category, the corresponding weight is larger. If the indicator is weakly associated with the category, the weight tends to zero. Using this matrix, we can convert the raw set of metrics into a vector of category representation:

$$H_i^t = C \cdot (D_i^t)^T \tag{8}$$

Among them, H_i^t represents the performance results of students on different dynamic assessment indicator categories. Compared with the single achievement, the category performance vector can more clearly present the student ability structure. For example, if a student's grammar accuracy is high, but the reasoning level is still insufficient after writing revision, his language basic indicators may be at a good level, while his cognitive processing indicators and learning regulation indicators still need to be strengthened. If students have insufficient oral fluency, but can quickly correct pronunciation and expression order after feedback, their current task performance is not high, but the interactive feedback index and potential development level are good.

In order to reduce the subjective bias of manual evaluation, natural language processing and learning behavior analysis methods are introduced into the index classification process. The features of the written text are extracted by vocabulary richness, grammatical error density, discourse cohesion word distribution and topic relevance. The spoken language data were analyzed by speech recognition results, pause position, speed stability and pronunciation similarity. Reading and listening data were classified by answer path, number of review, reaction time and error type. The feedback data were then coded according to whether the students took the advice, whether they completed the revision, and the revision quality change. In this way, the classification of dynamic assessment indicators not only has the calculation basis, but also can retain the teaching meaning of English learning activities.

2.2.2 Deep Learning Target Mapping methods

After completing the metric classification, it is also necessary to map the different metric categories to the deep learning objectives. The goals of deep learning in English teaching mainly include deep understanding, comprehensive expression, transfer application, critical thinking, collaborative communication and independent reflection. The value of dynamic assessment is not in generating more scores, but in revealing how well students are achieving these goals. To this end, we construct a mapping matrix between indicator categories and deep learning objectives:

$$M = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1p} \\ m_{21} & m_{22} & \cdots & m_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ m_{q1} & m_{q2} & \cdots & m_{qp} \end{bmatrix} \tag{9}$$

Here, p represents the number of deep learning objectives, and m_{rk} represents the support strength of the r-th dynamic evaluation index for the KTH deep learning objective. Taking writing as an example, grammatical accuracy mainly supports the basis of language expression, discourse structure and argument organization are closer to comprehensive expression and critical thinking, and revision records and feedback acceptance are closely related to independent reflection. Therefore, different data in the same learning task can be mapped to different targets, and the same deep learning target can also be supported by multiple types of indicators. The deep learning target adaptation result of student i in period t can be expressed as follows.

$$A_i^t = H_i^t \cdot M \tag{10}$$

$A_i^t = (a_{i1}^t, a_{i2}^t, \cdots, a_{ip}^t)$ represents the student's adaptation level on each deep learning objective. If the corresponding value of a certain target is low, it means that the evidence of ability under this target is insufficient, and teachers need to increase the corresponding teaching scaffold. For example, when deep understanding is insufficient, discourse structure analysis, keyword chain sorting and reasoning problem explanation can be added. When the comprehensive expression is insufficient, the training of paragraph organization, sentence group expansion and opinion support can be strengthened. When self-reflection is insufficient, students can be asked to submit modification instructions, error attribution sheets, and learning logs. In order to make the mapping results serve the teaching regulation, this paper further sets the target gap value:

$$\Delta_i^t = T_i - A_i^t \tag{11}$$

Here, T_i represents the target requirements set according to the course objectives, learning phases, and student bases, and Δ_i^t represents the gap between the current learning state and the target requirements. The larger the gap value, the stronger the teaching support required for this goal. Teachers can choose different intervention methods according to the type of gap. When the gap in language foundation is large, rule explanation and hierarchical practice are adopted. When the cognitive processing gap was large, problem chain, case comparison and task reconstruction were used. When the learning adjustment gap was large, intervention was carried out through learning schedule, feedback record sheet and phased reflection task.

Figure 3: Dynamic evaluation metric classification and deep learning target mapping paths

Figure 3 shows that there is a continuous transformation relationship between dynamic evaluation metric classification and deep learning target mapping. After index extraction, the data in the learning process first enter different evaluation categories, and then are transformed into deep learning goal attainment states through the mapping matrix. This path avoids the problem of disconnection between evaluation indicators and teaching objectives, and enables teachers to judge what kind of support students need from the data results, rather than only obtaining a total score. Through this method, the dynamic evaluation of English teaching can form a closed loop of "data collection, index classification, target mapping, gap diagnosis, teaching adjustment", and provide a clear input structure for the subsequent construction of teaching effectiveness optimization model.

3 Design of English teaching dynamic assessment adaptation mechanism based on deep learning objectives

3.1 Analysis of the Structure of English Learning Ability under the Goal Orientation of Deep Learning

The English learning ability structure under the goal-oriented of deep learning no longer understands students 'ability as the simple accumulation of vocabulary, grammar accuracy and test scores, but emphasizes the comprehensive level of students 'understanding, expression, transfer, evaluation and reflection in real language tasks. Dynamic assessment in English teaching needs to focus on this ability structure, so that the evaluation results can explain the questions of "can I use", "can I transfer" and "can I improve according to feedback". Therefore, the structure of English learning ability should include six parts: language knowledge foundation, discourse understanding ability, expression generation ability, thinking processing ability, interaction and collaboration ability and self-regulation ability, which support each other and form the basis for students to achieve deep learning goals.In the specific modeling, this paper expresses the English learning ability status of the ith student in the t learning cycle as follows.

$$Q_i^t = \{V_i^t, U_i^t, E_i^t, C_i^t, I_i^t, R_i^t\} \tag{12}$$

V_i^t represents language knowledge base, which mainly includes vocabulary, grammar, syntax and phonetic rules. U_i^t represents discourse comprehension ability, reflecting students' grasp of the main idea, logical relationship and implicit meaning of the text. E_i^t represents expression generation ability, which reflects the quality of students' language organization in writing, speaking and task output. C_i^t represents thinking processing ability, including reasoning judgment, viewpoint construction and critical analysis. I_i^t represents the ability of interaction and collaboration, and reflects the quality of students 'participation in class discussion, peer assessment and feedback response. R_i^t indicates self-regulation ability and reflects the learning plan, error revision and reflective improvement level. The capability structure can make dynamic evaluation shift from single skill detection to multi-dimensional capability diagnosis.

There is a correspondence between deep learning objectives and English learning ability. The language knowledge foundation provides support for understanding and expression, the discourse comprehension ability is related to the deep meaning construction, the expression generation ability directly affects the learning outcome output, the thinking processing ability determines the depth of language use, the interaction and collaboration ability reflects the sociality of language learning, and the self-regulation ability affects students 'absorption and further development of feedback. In order to describe the support degree of the ability structure to the deep learning goal, this paper constructs the comprehensive ability value:

$$S_i^t = \alpha_1 V_i^t + \alpha_2 U_i^t + \alpha_3 E_i^t + \alpha_4 C_i^t + \alpha_5 I_i^t + \alpha_6 R_i^t \tag{13}$$

Among them, S_i^t represents the comprehensive ability level of deep learning of students in the t-th cycle, α_1 to α_6 are the weights of different ability dimensions, and $\sum_{k=1}^{6} \alpha_k = 1$ is satisfied. The weights are not fixed and can be adjusted according to the teaching stage. For example, the foundation stage can appropriately improve the weight of language knowledge foundation and discourse comprehension ability. In the comprehensive task stage, the weights of expression generation, thinking processing and self-regulation ability should be increased. In this way, the dynamic assessment results can be consistent with the teaching objectives, instead of being led by a single test score.

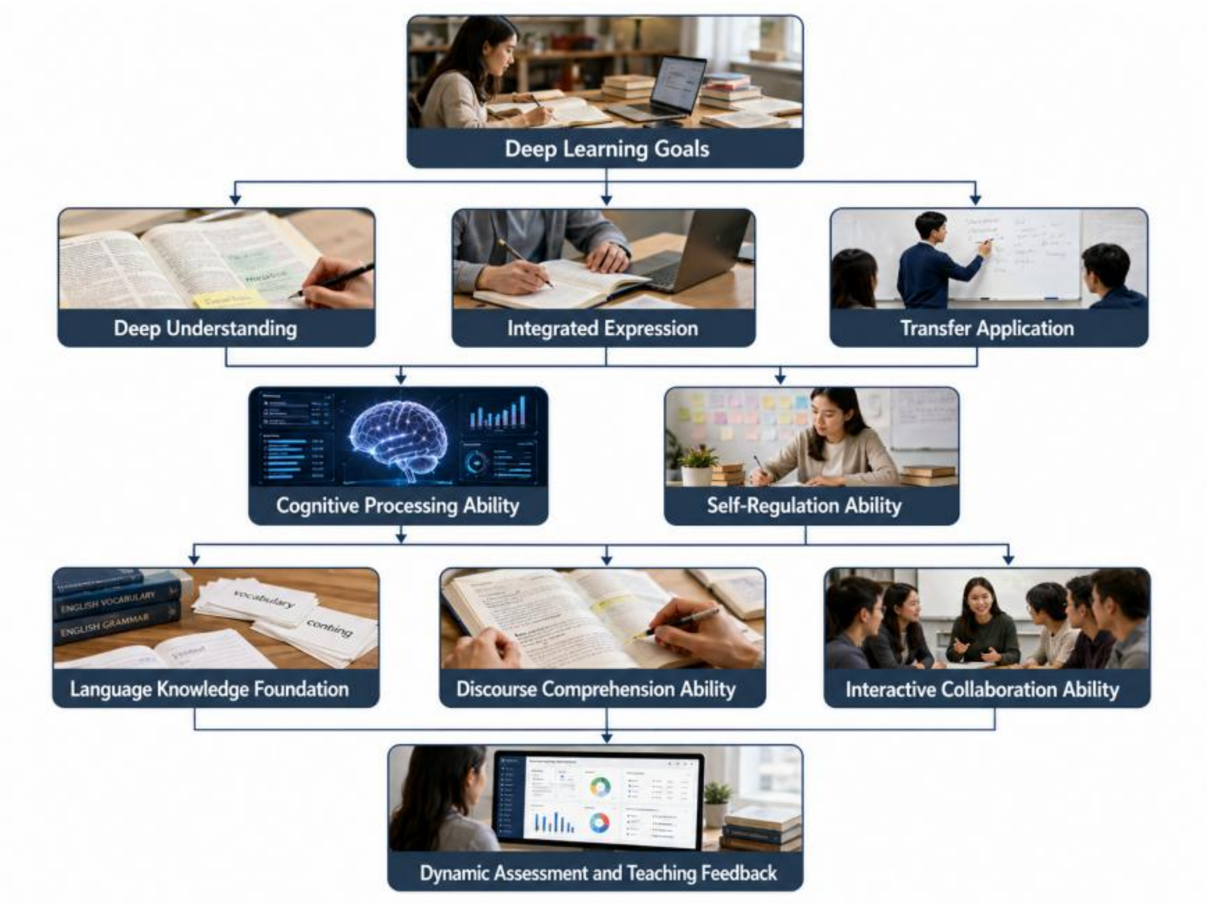

Figure 4: English learning competence structure under deep learning goal orientation

Figure 4 shows that the English learning ability structure is not linearly arranged, but forms a multi-dimensional support relationship around the deep learning goal. After collecting the data of the learning process, the dynamic evaluation system can identify the performance of students in different ability dimensions respectively, and then feed back the diagnosis results to the instructional design. For students with weak language knowledge, teachers can increase the training of lexical chunks, syntactic structures and phonetic rules. For students with insufficient comprehension ability, discourse logic analysis and problem chain guidance can be strengthened. For students with insufficient expression generation ability, exemplar breakdown, paragraph reconstruction and oral task output can be added. For students with insufficient self-regulation ability, it is necessary to strengthen their reflective consciousness through learning logs, revision instructions and feedback tracking. Thus, the deep learning objectives can be transformed into an English learning competence structure that can be analyzed, fed back and optimized.

3.2 Dynamic evaluation of English teaching and Teaching effectiveness optimization model construction

This paper constructs a dynamic evaluation and teaching effectiveness optimization model of English teaching based on deep learning objectives, which aims to integrate students 'learning process data, ability structure diagnosis results and teachers 'teaching regulation behaviors into a unified computing framework. Different from static teaching evaluation, this model does not take one-time performance as the main basis, but according to students' language performance, feedback response and ability gain in the learning cycle, judges the adaptation degree between teaching objectives and learning states, and generates the corresponding teaching optimization scheme.

3.2.1 Dynamic evaluation of adaptation principles

The construction of dynamic assessment model of English teaching should follow the following principles:

(1) Goal-oriented principle. The evaluation indicators should focus on the deep learning objectives such as deep understanding, comprehensive expression, transfer application, critical thinking and independent reflection, so as to avoid the evaluation content from deviating from the direction of curriculum training.

(2) Process diagnosis principle. The evaluation system should not only record the learning results, but also analyze the change process of students in task completion, feedback acceptance, error revision and strategy adjustment.

(3) Principle of data fusion. The model should integrate multi-source data such as text, speech, test, classroom interaction and platform log, so that the evaluation results can reflect the real state of English learning.

(4) Teaching adjustable principle. The evaluation results need to be able to be translated into instructional adjustment measures that can be implemented by teachers, such as task reorganization, hierarchical feedback, scaffolding prompts, and learning path correction.

3.2.2 Teaching effectiveness optimization index system

According to the above principles, this paper constructs the dynamic assessment of English teaching and teaching effectiveness optimization index system, as shown in Table 1.

Table 1: Dynamic assessment of English teaching and teaching effectiveness optimization index system

Primary Indicator	Secondary Indicator	Indicator Meaning
S1 Learning Process Engagement	S11 Task Completion Rate	The proportion of students who complete listening, speaking, reading, and writing tasks on time
	S12 Learning Continuity	Stability of learning duration, login frequency, and task intervals
	S13 Classroom Interaction Degree	Students' participation in questioning, discussion, peer assessment, and feedback response
S2 Language Performance Quality	S21 Lexical and Grammatical Accuracy	Use of vocabulary, syntactic structure, and control of grammatical errors
	S22 Discourse Organization Ability	Paragraph cohesion, logical development, and thematic consistency
	S23 Oral Expression Quality	Pronunciation clarity, speech-rate stability, and expression completeness
S3 Achievement of Deep Learning Objectives	S31 Deep Understanding	Understanding of text gist, logical relations, and implicit meanings
	S32 Transfer Application	Ability to apply language knowledge to new contexts and integrated tasks
	S33 Critical Expression	Ability in viewpoint judgment, evidence use, and argumentative development
S4 Teaching Feedback Effectiveness	S41 Feedback Adoption Rate	The extent to which students adopt suggestions from teachers or the system
	S42 Revision Gain	The improvement in learning performance before and after feedback
	S43 Teaching Adjustment Matching Degree	The degree of alignment between teachers' interventions and students' ability weaknesses

3.2.3 Dynamic evaluation of adaptation model design

Let the index vector of student i in the t-th learning cycle be as follows.

$$X_i^t = (x_{i1}^t, x_{i2}^t, \cdots, x_{in}^t) \tag{14}$$

Here, n represents the number of secondary indicators involved in the evaluation. In order to reflect the contribution of different indicators to the optimization of teaching effectiveness, the weight vector is set as follows.

$$W = (w_1, w_2, \cdots, w_n), \quad \sum_{j=1}^{n} w_j = 1 \tag{15}$$

Then the dynamic evaluation value of students in the current learning state can be expressed as follows.

$$D_i^t = \sum_{j=1}^{n} w_j\, x_{ij}^t \tag{16}$$

Only the current evaluation value is still not enough to judge the change of teaching effectiveness, so the model further introduces the ability gain term. Let D_i^{t+1} denote the student's evaluation value after teacher feedback and task revision, then the learning gain is as follows.

$$G_i^t = \frac{D_i^{t+1} - D_i^t}{D_i^t + \varepsilon} \tag{17}$$

Here, ε is the smoothing factor, which is used to avoid the denominator being zero. When G_i^t was high, it showed that teaching feedback had a strong promotion effect on students 'ability development. When D_i^t is high and G_i^t is low, it means that students' current performance is stable, but the teaching task may not be challenging enough. When D_i^t and G_i^t are both low, the system should prompt the teacher to increase the base support and hierarchical guidance.

In order to realize the optimization of teaching effectiveness, the goal gap, learning gain and feedback response are incorporated into the comprehensive optimization function:

$$O_i^t = \beta_1 D_i^t + \beta_2 G_i^t + \beta_3 R_i^t - \beta_4 \Delta_i^t \tag{18}$$

Here, O_i^t represents the teaching effectiveness optimization value of student i in period t, R_i^t represents the feedback response level, Δ_i^t represents the deep learning goal gap, and β_1 to β_4 are the regulation coefficients. This function can reflect learning results, development range, feedback absorption and goal gap at the same time, avoiding the model simply pursuing score improvement.

While the model is running, the system generates teaching suggestions based on changes in O_i^t. If the target gap is concentrated in the dimension of deep understanding, teachers can add discourse structure analysis and reasoning tasks. If the gap was concentrated in the comprehensive expression dimension, writing reconstruction, oral presentation and peer evaluation could be arranged. If self-regulation is insufficient, students' reflection is enhanced through learning logs, revision records and stage feedback sheets. Therefore, the dynamic evaluation results could enter the links of instructional design, classroom implementation and after-class feedback, forming a closed loop of teaching effectiveness optimization for deep learning objectives.

4 Research on the optimization effect of teaching efficiency

4.1 Experimental scheme design

In order to test the actual effect of the dynamic assessment adaptation mechanism of English teaching based on deep learning objectives on the optimization of teaching effectiveness, this paper adopted a quasi-experimental research method to compare and analyze the learning performance, feedback response and the achievement of deep learning goals before and after the application of the model. The experimental design adheres to the control principles of the same teaching material, the same teaching progress, similar learning foundation and the same evaluation cycle, and tries to reduce the influence of teachers' differences, course content differences and learning foundation differences on the experimental results, so that the optimization effect of the model can be verified in a more stable teaching environment.

(1) Research object

In this study, two parallel college English teaching classes for 2024 non-English majors in a local undergraduate university were selected as the research objects, with a total of 96 students. Among them, there were 48 students in the experimental class and 48 in the control class. Before the experiment, the students of the two classes were given a pre-test of comprehensive English ability, which included four parts: reading comprehension, writing expression, oral task and vocabulary and grammar application, with a full score of 100. The mean pre-test score of the experimental class was 71.36, and the mean pre-test score of the control class was 70.92. The independent sample test results showed that the difference between the two classes was not significant, indicating that the two groups of students were comparable before the experiment. The experimental class adopted the dynamic assessment and deep learning target adaptation model constructed in this paper, and the control class adopted the conventional classroom evaluation method, that is, the unit test, assignment scoring and teacher classroom observation as the main basis.

(2) Experiment cycle and teaching arrangement

The experimental period was set to 16 weeks, with 4 periods per week, for a total of 64 periods. Both classes use the same comprehensive college English course, which covers thematic reading, topic discussion, writing exercises, listening and speaking tasks and stage tests. At the end of each teaching unit, the experimental class collects the data of students 'learning process by the system, and generates dynamic assessment results according to six dimensions: language foundation, discourse understanding, expression generation, thinking processing, interaction and collaboration, and self-regulation. Teachers adjusted subsequent teaching tasks based on the system diagnosis results, including stratified reading materials, writing modification checklist, oral expression scaffold, peer assessment task, and learning reflection recording. The control class maintains the regular teaching process, and the teachers make comments according to the homework performance and test results.

(3) Data collection content

The experimental data consists of learning platform logs, classroom interaction records, written texts, spoken audio, reading tests, and questionnaire results. During the experiment, a total of 28640 effective learning behavior logs, 768 writing texts, 1536 oral audio segments, 2304 reading test answer records, 1248 classroom interaction records, and 192 student feedback questionnaires were collected. The grammatical error rate, vocabulary richness, discourse coherence and revision range of the writing text were extracted by the automatic writing evaluation module. Speech recognition module was used to extract articulation clarity, pause frequency and expression integrity of spoken audio. Platform logs are used to calculate learning persistence, task completion rate, and feedback adoption rate. In order to ensure the validity of the data, the samples with a missing rate of more than 20% were not included in the final statistics, and the abnormal learning time and repeated submission records were corrected after verification of the platform log.

Table 2: Experimental protocol and data acquisition arrangement

Item	Experimental Class	Control Class
Number of students	48 students	48 students
Experimental period	16 weeks	16 weeks
Teaching content	College English integrated course, reading, writing, listening and speaking tasks	College English integrated course, reading, writing, listening and speaking tasks
Evaluation method	Dynamic assessment, objective mapping, feedback tracking, learning gain analysis	Unit tests, assignment scoring, classroom observation
Data sources	Platform logs, writing texts, oral audio, reading records, interaction feedback	Test scores, assignment scores, teacher records
Main indicators	Deep understanding, comprehensive expression, transfer application, feedback adoption rate, revision gain	Reading score, writing score, oral score, overall score
Analysis tools	Python 3.11, SPSS 26.0, learning analytics module	SPSS 26.0

(4) Research methods

In this paper, quantitative analysis and process analysis are combined. The quantitative part mainly compared the differences between the experimental class and the control class in the pre-test, post-test, learning gain, writing quality, oral performance and deep learning goal achievement. The process analysis mainly examined the revision behavior, task completion stability and autonomous reflection quality of students in the experimental class after dynamic feedback. In the statistical processing, SPSS 26.0 was used to complete the independent sample test and paired sample test, and Python 3.11 was used to clean and calculate the learning log, text features and spoken features. The significance level was set at 0.05. Through the above experimental scheme, the effect of the proposed model on the optimization of English teaching effectiveness can be tested from three levels: learning results, learning process and teaching feedback.

4.2 Results Analysis

Table 3 shows the changes in the main learning efficacy indicators of the experimental class and the control class before and after the experiment. In order to ensure that the analysis results can reflect the role of dynamic evaluation and deep learning target adaptation mechanism, this paper did not use the final score alone as the judgment basis, but included deep understanding, comprehensive expression, transfer application, feedback adoption rate, revision gain and comprehensive post-test scores into the comparison scope. All indicators were converted into a hundred-point system, where feedback adoption rate indicated the proportion of students effectively using teacher feedback, system prompts and peer suggestions, and revision gain indicated the improvement of students 'performance on text, speaking or reading tasks after receiving feedback.

Table 3: Changes in teaching effectiveness indicators of the two groups of students before and after the experiment

Indicator	Experimental Class Pre-test	Experimental Class Post-test	Improvement in Experimental Class	Control Class Pre-test	Control Class Post-test	Improvement in Control Class
Deep Understanding	70.84	82.67	11.83	70.52	75.48	4.96
Comprehensive Expression	69.36	81.92	12.56	69.11	74.63	5.52
Transfer Application	67.95	79.84	11.89	68.24	72.91	4.67
Critical Expression	66.48	77.53	11.05	66.71	71.26	4.55
Feedback Adoption Rate	61.72	84.36	22.64	62.18	70.43	8.25
Revision Gain	58.46	80.27	21.81	59.03	68.52	9.49
Comprehensive Post-test Score	71.36	84.21	12.85	70.92	76.08	5.16

As can be seen from Table 3, the improvement of the experimental class in all indicators is higher than that of the control class. The feedback acceptance rate increased from 61.72 to 84.36, an increase of 22.64 percentage points; The revision gain is increased from 58.46 to 80.27, an increase of 21.81 percentage points. This indicates that the dynamic assessment mechanism can prompt students to understand and use feedback more effectively. Although there was a certain improvement in the control class, the improvement was mainly concentrated in the routine performance and basic task completion, and the changes in transfer application, critical expression and revision gain were relatively limited. It can be seen that unit tests and homework reviews can promote students to complete knowledge consolidation, but they do not support the ability diagnosis and feedback redevelopment in the learning process.

Figure 5 shows the change trend of comprehensive learning performance between experimental and control classes. In this paper, the pre-test, week 8, week 12 and week 16 were used as observation nodes to calculate the comprehensive scores of students on four indicators: language performance, task completion, deep learning goal achievement and feedback response.

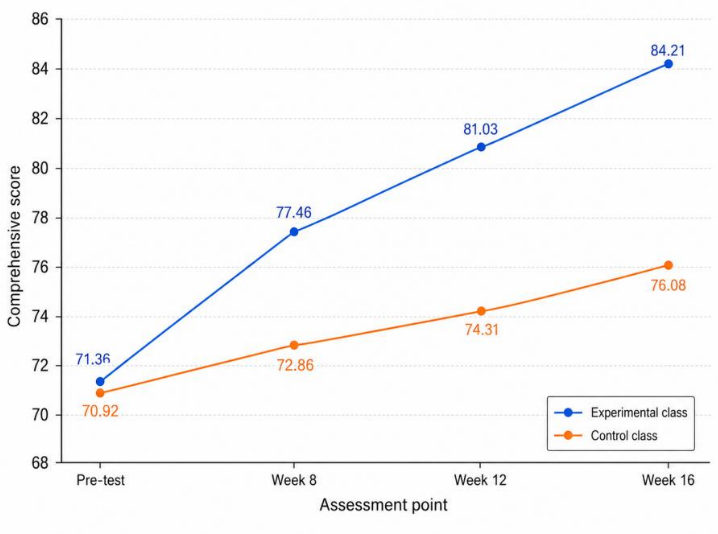

Figure 5: Change trend of comprehensive learning performance between experimental class and control class

It can be seen from Figure 5 that the comprehensive learning performance of the experimental class increased significantly from the 8th week, and the growth trend remained stable after the 12th week. The reason is that the dynamic assessment model does not wait for unified evaluation at the end of the semester, but generates learning state portraits after each teaching unit, so that teachers can timely discover students 'deficiencies in discourse comprehension, expression generation and self-regulation, and intervene through hierarchical tasks. The comprehensive score of the control class also showed an upward trend, but the growth rate was slow, indicating that the conventional teaching evaluation had a certain role in promoting students 'phased learning, but it was difficult to continuously track the change of ability after feedback.

In order to further observe the differences in the attainment of deep learning goals, this paper compared the five indicators of deep understanding, comprehensive expression, transfer application, critical expression and self-reflection and revision gain of the two groups of students after the experiment, and the results are shown in Figure 6.

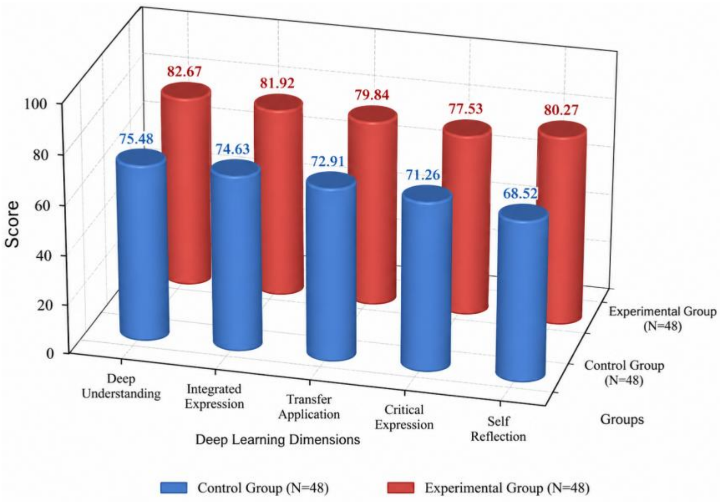

Figure 6: Comparison of the attainment of deep learning goals between the two groups of students after the experiment

Figure 6 shows that the experimental class scored higher than the control class on the five deep learning objectives. Among them, the difference of self-reflection was the most obvious, which was 80.27 in the experimental class and 68.52 in the control class, with a difference of 11.75. This result showed that the dynamic assessment and deep learning goal adaptation mechanism had a strong role in promoting students 'learning reflection and feedback revision. Students in the experimental class need to check the system diagnosis results after each writing, speaking and reading task, and complete the revision instructions combined with teacher feedback, which makes students gradually form the awareness of error recognition, cause analysis and strategy adjustment. In the control class, results were mainly used. Although students were able to obtain correct answers and scoring results, their understanding of their own ability shortcomings was not specific enough, and the depth of feedback use was relatively insufficient.

From the perspective of comprehensive expression and transfer application indicators, the post-test scores of the experimental class were 81.92 and 79.84, respectively, which were significantly higher than those of the control class. This shows that the dynamic evaluation mechanism can transform learning data into specific teaching support, so that students no longer stop at the local correction of vocabulary and syntax level, but gradually pay attention to paragraph logic, idea expansion, context adaptation and expression purpose. Especially in the topic writing and oral discussion tasks, the system generated personalized feedback according to students' previous performance, and teachers arranged case breakdown, problem chain discussion and group mutual evaluation according to the common problems in the class to promote students to apply language knowledge to new communication tasks.

(2) Application effect analysis of teaching efficiency optimization model

In order to verify the application value of the proposed model in terms of evaluation efficiency and evaluation reliability, this paper compares it with the conventional manual evaluation method and a single automatic scoring method. Conventional manual evaluation methods mainly rely on teacher's correction, classroom observation and unit performance statistics. The single automatic scoring method mainly relies on the automatic scoring results of writing or speaking. The proposed model integrates learning process logs, language output features, feedback response records, and deep learning target mapping results. See Figure 7 for the average evaluation time of the three methods with the same data size.

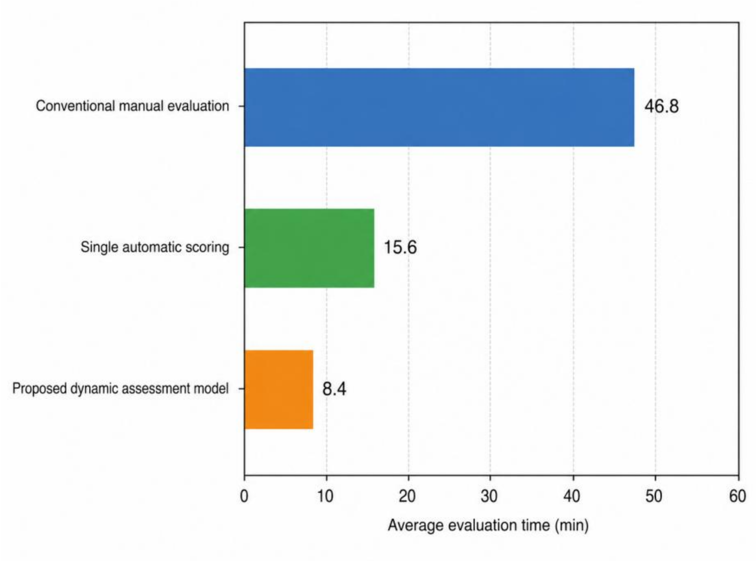

Figure 7: Comparison of average evaluation time of different evaluation methods

It can be seen from Figure 7 that the conventional manual evaluation method takes an average of 46.8 minutes, and the main time consumption is concentrated in the writing correction, oral performance judgment, and feedback consolidation. The single automatic scoring method shortened the evaluation time to 15.6 minutes, but its evaluation content focused on the surface features of text or speech, and the coverage of students 'feedback adoption, task revision and deep learning goal achievement was insufficient. The average evaluation time of the dynamic evaluation model in this paper is 8.4 minutes, which is significantly lower than the two comparison methods. This is because the model has completed log recording, text feature extraction, speech recognition and indicator standardization in the data collection stage, so that teachers can directly view the learning performance portrait and target gap results, and then make teaching adjustments for typical problems.

In terms of evaluation error, this paper uses the teacher review results as a reference to compare the comprehensive evaluation results output by the three methods with the review results. The average error of the conventional manual evaluation method is 7.82%, the average error of the single automatic scoring method is 6.37%, and the average error of the proposed model is 3.18%. Conventional manual evaluation is easily affected by teachers' subjective experience, grading time and class size. Although a single automatic grading can improve consistency, it is not sufficient to identify context expression, thinking quality and feedback changes. The model reduces the deviation of single index through multi-source data fusion and target mapping mechanism, making the evaluation results closer to the real learning state of students.

Further analysis of the learning process records of the students in the experimental class showed that the dynamic assessment mechanism had a more obvious promotion effect on the middle level students. Before the experiment, there were 23 students in the experimental class with comprehensive scores between 70 and 79, accounting for 47.92%. After the experiment, the number of students with this score dropped to 13, while the number of students with more than 80 scores increased from 9 to 28. This shows that the model can not only improve the comprehensive output quality of high-level students, but also help medium-level students enter a higher performance level through feedback revision. The number of students in low sections was reduced from 7 to 2, indicating that the systematic scaffold had a certain compensation effect on students with weak foundation.

From the perspective of teaching feedback, the teachers of the experimental class made two significant adjustments to the teaching content after the eighth week. One is to increase the classroom training of "topic sentence recognition-logical connective analysis-implicit meaning judgment" to solve the problem of high error rate of reading reasoning. The other is to solve the problem of weak writing argumentation, adding the task chain of "opinion raising - evidence selection - paragraph expansion - peer review". After two adjustments, the reading deep comprehension score of the experimental class increased from 76.15 in the 8th week to 82.67 in the 16th week, and the writing comprehensive expression score increased from 75.82 to 81.92. This result shows that the goal gap of the model output can not only describe the student problem, but also support the teacher to carry out targeted teaching reconstruction.

5 Conclusion

Focusing on the adaptation between dynamic assessment and deep learning objectives in English teaching, this paper constructs a dynamic assessment model for teaching effectiveness optimization. Based on the English learning process data, this study integrates platform behavior logs, writing texts, oral audio, reading tests, classroom interaction and feedback records into a unified analysis framework, and describes the structure of students 'learning ability from the dimensions of language knowledge foundation, discourse comprehension, expression generation, thinking processing, interaction and collaboration, and self-regulation. On this basis, this paper designs a dynamic evaluation index classification method and a deep learning target mapping mechanism, so that the evaluation results can shift from single performance judgment to ability diagnosis, target gap identification and teaching feedback optimization. The experimental results show that after using this model, the experimental class has higher improvement in indicators such as deep understanding, comprehensive expression, transfer application, critical expression and independent reflection than the control class, and the feedback acceptance rate and revision gain also show significant improvement. At the same time, the evaluation time and evaluation error of the model are better than that of conventional manual evaluation and single automatic scoring method, indicating that multi-source data fusion and target mapping can improve the efficiency and reliability of English teaching evaluation. This study provided an operational path for dynamic assessment, personalized feedback and teaching strategy adjustment in English classroom, and also provided a method reference for the optimization of English teaching effectiveness supported by artificial intelligence technology.

References

[1] Kushki A, Nassaji H, Rahimi M. Interventionist and interactionist dynamic assessment of argumentative writing in an EFL program[J]. System, 2022, 107: 102800.

[2] Kushki A, Rahimi M, Davin K J. Dynamic assessment of argumentative writing: Mediating task response[J]. Assessing Writing, 2022, 52: 100606.

[3] Rassaei E. Implementing mobile-mediated dynamic assessment for teaching request forms to EFL learners[J]. Computer Assisted Language Learning, 2023, 36(3): 257-287.

[4] Kafipour R, Khoshnood A. Effect of feedback through dynamic assessment on EFL field-dependent and field-independent learners' speaking skill development[C]// Frontiers in Education. Frontiers Media SA, 2023, 8: 1049680.

[5] Zadkhast M, Rezvani E, Lotfi A R. Effects of concurrent and cumulative group dynamic assessments on EFL learners' development of reading comprehension micro-skills[J]. Language Testing in Asia, 2023, 13(1): 29.

[6] Randall T S, Urbanski K. Development of a computerized dynamic assessment program for second language grammar instruction and assessment[J]. Language and Sociocultural Theory, 2023, 10(1): 50-81.

[7] Anam S, Akhiriyah S, Iswati H D. Looking into the role of dynamic assessment in English grammar mastery of Indonesian EFL learners[C]//Unima International Conference on Social Sciences and Humanities (UNICSSH 2022). Atlantis Press, 2023: 571-578.

[8] Kumar A, Rupley W, McKeown D, et al. Beyond the red pen: Using dynamic assessment to mediate writing mechanics issues among ESL learners[J]. Journal of Contemporary Language Research, 2023, 2(4): 171-180.

[9] Masrul M, Rasyidah U, Yuliani S, et al. The Implementation of Dynamic Assessment in EFL Learners' Writing[J]. World Journal of English Language, 2023, 13(5): 191-199.

[10] Suri S S. Teachers' strategies to enhance deeper learning skills in English language classes[J]. International Journal of Linguistics, Literature and Translation, 2024, 7(3): 118-125.

[11] Peña-Acuña B, Corga Fernandes Durão R. Learning English as a second language with artificial intelligence for prospective teachers: a systematic review[C]//Frontiers in education. Frontiers Media SA, 2024, 9: 1490067.

[12] Kristiawan D, Bashar K, Pradana D A. Artificial intelligence in English language learning: A systematic review of AI tools, applications, and pedagogical outcomes[J]. The Art of Teaching English as a Foreign Language (TATEFL), 2024, 5(2): 207-218.

[13] Moorhouse B L. Generative artificial intelligence and ELT[J]. ELT Journal, 2024, 78(4): 378-392.

[14] Tolstykh O M, Oshchepkova T. Beyond ChatGPT: Roles that artificial intelligence tools can play in an English language classroom[J]. Discover Artificial Intelligence, 2024, 4(1): 60.

[15] Al-khresheh M H. The Future of Artificial Intelligence in English Language Teaching: Pros and Cons of ChatGPT Implementation through a Systematic Review[J]. Language Teaching Research Quarterly, 2024, 43: 54-80.

[16] Ulla M B, Advincula M J C, Mombay C D S, et al. How can GenAI foster an inclusive language classroom? A critical language pedagogy perspective from Philippine university teachers[J]. Computers and Education: Artificial Intelligence, 2024, 7: 100314.

[17] Aldosemani T I, Assalahi H, Lhothali A, et al. Automated writing evaluation in EFL contexts: A review of effectiveness, impact, and pedagogical implications[J]. International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT), 2023, 13(1): 1-19.

[18] Sari E, Han T. The impact of automated writing evaluation on English as a foreign language learners' writing self-efficacy, self-regulation, anxiety, and performance[J]. Journal of Computer Assisted Learning, 2024, 40(5): 2065-2080.

[19] Dennis N K. Using AI-Powered Speech Recognition Technology to Improve English Pronunciation and Speaking Skills[J]. IAFOR Journal of Education, 2024, 12(2): 107-126.

[20] Sajja R, Sermet Y, Cikmaz M, et al. Artificial intelligence-enabled intelligent assistant for personalized and adaptive learning in higher education[J]. Information, 2024, 15(10): 596.