

Implementation Pathways and Effectiveness Verification of Formative Assessment in College English Teaching

Yuqing Zhang^{1,2}, Zhiyuan Jin³

¹School of Graduates, Xi'an International Studies University, Xi'an 710128, Shaanxi, China

²School of Culture and Education, Shaanxi University of Science & Technology, Xi'an 710021, Shaanxi, China

³School of Education, Xi'an International Studies University, Xi'an 710128, Shaanxi, China

Corresponding author: Yuqing Zhang (e-mail: 18091063830@163.com)

ABSTRACT In the context of high-quality development in higher education and the continuous reform of college English teaching, traditional summative assessment is no longer sufficient to evaluate language application, autonomous learning, and technical communication competencies required in engineering fields such as electromagnetic waves, antennas, and propagation. Formative assessment, with its process-oriented, developmental, and interactive characteristics, emphasizes continuous tracking, feedback, and improvement throughout learning, and can support the transition from knowledge transmission to ability cultivation. This study constructs an implementation pathway for formative assessment in college English teaching and verifies its effectiveness through a one-academic-year quasi-experiment. Two comparable undergraduate classes were selected as experimental and control groups, with the experimental group adopting a formative assessment system integrating goal clarification, learning evidence collection, multi-source feedback, self- and peer-assessment, and improvement-oriented re-evaluation. Academic performance, learning motivation, autonomous learning ability, classroom participation, and language application ability were analyzed using quantitative tests and interviews. The results indicate that formative assessment significantly improves students' oral expression, writing logic, cooperative communication, and learning initiative compared with traditional assessment. The study provides an operational model for engineering-oriented English teaching and offers support for cultivating technical communication abilities in advanced electromagnetic and related disciplines.

INDEX TERMS formative assessment, college english, engineering communication, antenna and propagation, learning analytics

I. INTRODUCTION

A. RESEARCH BACKGROUND

AS China's higher education enters a new stage of connotative development, talent cultivation has transitioned from scale expansion to quality improvement, necessitating advanced international communication and innovative practice abilities to meet the evolving technical standards of the industrial sector. This shift requires students to master cross-cultural collaboration and autonomous learning to effectively drive the modernization of technologies and high-end industrial manufacturing. College English, as a fundamental course in higher education, plays an important role in enhancing students' language application abilities, cultivating international perspectives and comprehensive literacy. Its teaching quality directly affects the effectiveness of cultivating high-quality composite talents [1-3].

B. RESEARCH SIGNIFICANCE

1) Theoretical Significance

From the perspective of educational evaluation theory, the core value of formative assessment lies in transitioning from the traditional measurement screening paradigm toward a developmental logic that integrates student-centered learning with the technical complexities of industrial research. This approach constructs an evaluation framework capable of fostering the analytical skills and professional literacy required to innovate within the field of science and modern industrial applications. The theoretical contribution of this article is mainly reflected in three aspects:

Enrich the application research of formative assessment in the field of college English teaching, improve the theoretical framework of "evaluation teaching learning" trinity, and provide empirical support for the theory of foreign language education evaluation. By integrating constructivism, humanism, and closed-loop teaching theory, we aim to clarify the mechanism of formative assessment in language learning scenarios and address the gap between theory and practice

in existing research [4].

2) Practical Significance

The normalized implementation of formative assessment is a practical path to solve the current difficulties in college English teaching, and its practical value is reflected in four aspects:

1. Provide a feasible, replicable, and promotable formative evaluation operation plan for college English teachers to solve the practical problem of "wanting to do but not being able to do it, and not doing it systematically". By implementing standardized guidelines, evaluation scales, and closed-loop process design, we aim to lower the implementation threshold for teachers and avoid fragmented operations [5].

2. Effectively stimulate students' learning motivation, enhance their self-learning ability and classroom participation, promote the transformation of language knowledge into language application ability, and effectively improve teaching effectiveness. The empirical results show that formative assessment can significantly improve students' oral expression, writing logic, and cooperative communication abilities, and alleviate the phenomenon of "high scores but low abilities" and "mute English" [6].

3. Help teachers grasp learning dynamics in real time, accurately diagnose teaching problems, optimize teaching strategies, and achieve the goal of "promoting learning, teaching, and improvement through evaluation". Through process data collection and feedback, teachers can adjust teaching focus and difficulty points in a timely manner, optimize teaching design, and improve teaching accuracy and efficiency.

4. Promote the transformation of college English teaching evaluation from "single score" to "multiple abilities", and help improve the quality of talent cultivation in universities. Raise the weight of formative assessment to 50%, highlighting the focus on the learning process and ability development, in line with the requirements of the new era's talent cultivation for the trinity of "knowledge, ability, and value" [7].

C. CURRENT RESEARCH STATUS AT HOME AND ABROAD

Domestic scholars have conducted extensive research on the application of formative assessment in college English teaching, mainly focusing on three aspects:

1. Theoretical Introduction: Explain the concept, characteristics, value, and theoretical basis of formative assessment. For example, Zhou Xuelin (2011) systematically introduced the connotation of formative assessment and its application value in college English teaching.

2. Mode construction: Propose a combination of classroom evaluation, homework evaluation, online evaluation, project evaluation, etc. For example, Wang Chuming (2014) proposed the formative evaluation mode of "writing promotes learning", emphasizing the improvement of language ability through writing feedback.

However, there are still significant shortcomings in existing research:

1. The implementation path is fragmented and superficial, mostly staying at the level of "increasing normal scores", without forming a closed-loop system, and lacking the design of the entire process of "evaluation feedback improvement".

2. The evaluation method is single, lacking in-depth design of key aspects such as self-evaluation, peer evaluation, feedback, and improvement. The process records are incomplete, the evidence is insufficient, and the standards are unclear.

3. The empirical research sample is small, the cycle is short, the effectiveness verification is not comprehensive enough, and the mechanism analysis is insufficient. Most studies only focus on academic performance, and the verification of core competencies such as learning motivation and self-directed learning ability is insufficient.

4. There is insufficient research on the adaptability of large class teaching, blended online and offline teaching, and students of different levels, which makes it difficult to solve the practical bottleneck of teachers' heavy workload.

Based on this, this article is based on the real teaching scenarios in Chinese universities, constructs a systematic implementation path, and conducts rigorous empirical research, which has strong theoretical and practical value.

The innovation lies in constructing a replicable "five-in-one" closed-loop path for large-class teaching, quantifying implementation effects, and solving the heavy workload bottleneck, rather than merely proving the effectiveness of feedback.

D. RESEARCH IDEAS, CONTENT, AND METHODS

1) Research Approach

Following the technical route of "theoretical construction dilemma analysis path design empirical verification strategy optimization":

1. Sort out the relevant theories of formative assessment, foreign language teaching, and educational psychology, define core concepts, and clarify the operational logic of formative assessment.

2. Analyze the current difficulties and causes in the evaluation of college English teaching, and reveal the limitations of traditional evaluation models.

3. Build a five in one formative evaluation implementation path of "goal implementation monitoring feedback improvement", covering the entire process of pre class, in class, and post class.

4. Select 126 undergraduate students from two natural classes at a certain university and conduct a quasi experimental study for one academic year to verify the implementation effect from five dimensions: academic performance, learning motivation, self-directed learning ability, classroom participation, and language application ability.

5. Based on empirical results and practical feedback, propose optimization strategies from the aspects of concep-

tual updates, institutional guarantees, technical support, and teacher professional development to provide guarantees for the normalized implementation of formative evaluation.

2) Research Content

1. Theoretical basis and core connotation of formative assessment: Sort out relevant theories, define the concept, characteristics, and value functions of formative assessment.

2. The realistic dilemma of college English teaching evaluation: Analyze the problems and causes in current evaluation concepts, content, methods, feedback, implementation, and other aspects.

3. Implementation path for formative assessment of college English: Construct an integrated implementation system of pre class, in class, and post class, design evaluation processes, tools, and weight allocation.

4. Verification of the effectiveness of formative assessment: Through quasi experimental research, quantitatively analyze the impact of formative assessment on academic performance, learning motivation, self-directed learning ability, classroom participation, and language application ability.

5. Optimization strategies for implementing formative evaluation: Provide actionable optimization suggestions from aspects such as philosophy, system, technology, teachers, feedback, etc.

3) Research Methods

1. Select two natural classes in the first year of non English major undergraduate studies at a certain university, and set up an experimental group (formative evaluation) and a control group (traditional evaluation) for a one-year comparative study. Control variables such as textbooks, teachers, and class hours to ensure the reliability of the experimental results.

2. Interview method: Conduct semi-structured interviews with 20 students and 5 teachers to obtain qualitative feedback and conduct in-depth analysis of the implementation experience and mechanism of formative evaluation.

3. Statistical analysis method: Use SPSS 26.0 to perform descriptive statistics, independent sample t-test, paired sample t-test, and correlation analysis on academic performance and questionnaire data to verify the effectiveness and significance of formative evaluation [8].

II. CORE CONCEPTS AND THEORETICAL FOUNDATIONS

A. DEFINITION OF CORE CONCEPTS

1) Formative Evaluation

Formative Assessment was first proposed by Scriven in 1967 and later developed by Black&William (1998). It refers to an evaluation method that is continuously carried out in the teaching process, aimed at diagnosing problems, providing feedback, improving teaching, and promoting development. It does not aim for ranking or screening, but focuses on

learning progress, effort level, improvement range, and ability enhancement, emphasizing the subject status of students in evaluation, encouraging self-evaluation, peer evaluation, and reflection [9-11]. Its core features can be summarized into five points:

Process oriented: Throughout the entire teaching process, from pre class preparation to post class reflection, rather than just at the end;

Developmental: with the fundamental goal of promoting students' comprehensive development and improving teachers' teaching, weakening the functions of screening and selection;

Diversity: Multiple evaluation subjects (teacher evaluation, self-evaluation, peer evaluation), content (knowledge, ability, literacy), and methods (observation, testing, homework, projects);

Interactivity: Emphasizing feedback and dialogue between teachers, students, and students, achieving cognitive upgrading and ability enhancement through interaction;

Closed loop: Form a cycle of "evaluation feedback improvement re evaluation" to ensure that the evaluation results can directly affect the optimization of teaching and learning.

Effect verification is the use of quantitative and qualitative methods to examine the true role of formative assessment in improving academic performance, learning motivation, self-directed learning ability, classroom participation, language application ability, and other aspects, revealing its effectiveness and mechanism of action, and providing empirical support for practical promotion [12].

B. THEORETICAL BASIS

Formative evaluation weakens the functions of identification and selection, strengthens motivation and support, protects students' confidence in learning through personalized feedback and positive encouragement, stimulates internal learning motivation, and embodies the humanistic educational philosophy of "putting student development at the center" [13,14].

Black&William (1998) pointed out in the classic study "Inside the BlackBox" that effective formative assessment includes five key elements:

1. Clear learning objectives and success criteria;
2. Rich collection of learning evidence;
3. Timely and effective feedback;
4. Student self-evaluation and peer evaluation;
5. Adjust teaching based on evaluation results. This theory provides a core framework for constructing the implementation path of formative evaluation in this article, ensuring that the evaluation process has the core functions of diagnosis, feedback, and improvement [15].

Hypotheses are derived from constructivism, feedback intervention theory, and self-regulated learning theory, instead of direct deduction from the definition of formative assessment.

III. THE REALISTIC DILEMMA OF COLLEGE ENGLISH TEACHING EVALUATION

A. EVALUATION PHILOSOPHY: EMPHASIZE TERMINATION OVER PROCESS, PRIORITIZE SCREENING OVER DEVELOPMENT

At present, there is a common deviation in the concept of "screening orientation" in the evaluation of English teaching in Chinese universities, which regards evaluation as a tool for dividing student levels and screening outstanding talents, rather than a means to promote learning. This concept is manifested in practice as: The proportion of final exam scores for summative exams is as high as 70% -80%, while formative evaluations are only used as "embellishments", accounting for 20% -30% and being mere formalities;

Teachers focus on "how to score and rank", while students focus on "how to improve scores and pass exams", which seriously weakens the core values of promoting learning and teaching through evaluation; The evaluation results are used for excellence and scholarship assessment, which strengthens the exam oriented thinking of "score first" and suppresses students' interest and initiative in learning [16]. This concept is seriously disconnected from the needs of talent cultivation in the new era and is difficult to adapt to the requirements of the connotative development of higher education.

B. EVALUATION CONTENT: EMPHASIZE KNOWLEDGE OVER ABILITY, FOCUS ON MEMORY OVER APPLICATION

The singularity of evaluation content is a prominent problem in current college English evaluation, mainly manifested as:

The evaluation focuses on written knowledge such as vocabulary, grammar, reading, and translation, with insufficient attention paid to core competencies such as oral expression, listening application, writing logic, collaborative communication, and critical thinking and innovation;

The exam question types are mainly objective questions (multiple-choice questions, fill in the blank questions), focusing on mechanical memory and surface understanding, and the assessment of language output ability (speaking, writing) is relatively low; The evaluation content is disconnected from the actual language application scenarios, and the phenomenon of "high scores but low abilities" and "mute English" among students is still common, which is difficult to meet the needs of the workplace and international communication [17].

This content orientation leads students to "learn for exams and memorize for scores", neglecting language accumulation and ability training, and unable to achieve the transformation of language knowledge into language application ability [18].

C. EVALUATION METHODS: SINGULARITY, SUPERFICIALITY, FORMALIZATION

The current formative assessment of college English is mostly a simple combination of "atten-

dance+homework+classroom answers+midterm exams", which has obvious problems of singularity and superficiality;

Single evaluation subject: mainly based on one-way evaluation by teachers, lacking the participation of multiple subjects such as student self-evaluation, peer evaluation, and parental evaluation;

The evaluation method is single: mainly based on paper and pencil testing and homework grading, lacking in-depth forms such as classroom observation, oral presentation, project tasks, and portfolio evaluation;

Incomplete process recording: only recording scores or grades, lacking tracking and recording of learning process, effort level, and progress;

Unclear standards: The evaluation criteria are vague and subjective, lacking a unified evaluation scale and operational norms, resulting in low credibility of the evaluation results.

This evaluation method cannot truly reflect students' learning process and ability development, and it is difficult to play the diagnostic and improvement functions of formative evaluation [19,20].

D. FEEDBACK MECHANISM: LAG, AMBIGUITY, ONE-WAY, LACK OF IMPROVEMENT GUIDANCE

Feedback is the soul link of formative evaluation and also the weakest part in current implementation. The feedback from most teachers has the following issues:

Feedback lag: In class task feedback is not timely, and the feedback cycle for homework after class is as long as 3-5 days, leading to problem accumulation;

Fuzzy feedback: Only providing "scores or grades" without specific and actionable improvement suggestions, such as vague evaluations like "very good" or "not bad", unable to convey effective information;

One way feedback: mainly one-way output from teachers to students, lacking teacher-student dialogue, peer feedback, and student self-evaluation, students are unable to actively participate in evaluation and reflection;

Lack of improvement loop: A "correction review" mechanism has not been established after feedback, and students do not know how to improve, resulting in feedback becoming mere formality.

This inefficient feedback mechanism renders the diagnostic and improvement functions of formative evaluation completely ineffective, making it impossible to achieve the goal of "promoting improvement through evaluation".

E. IMPLEMENTATION DIFFICULTIES: LARGE CLASS TEACHING, HEAVY WORKLOAD, LACK OF SYSTEM DESIGN

College English is mostly taught in large classes (50-100 people), and the heavy burden on teachers is a practical bottleneck that makes it difficult to implement formative assessment on a regular basis. The workload of manual recording, correction, and statistics is heavy, making it

difficult for teachers to continuously track and provide personalized feedback to each student;

The lack of unified implementation standards, evaluation tools, and technical support has led to the inability to conduct formative evaluations in a long-term or in-depth manner;

Under large class teaching, evaluation methods such as classroom observation and oral presentation are difficult to cover all students, leading to the phenomenon of "a few people participating and the majority watching";

The disconnect between evaluation and teaching has led many teachers to view evaluation as an "additional task", and the evaluation results have not been used to adjust teaching progress or optimize teaching methods, resulting in the formation of "two skins of evaluation".

These practical difficulties seriously constrain the implementation and effectiveness of formative evaluation.

IV. CONSTRUCTION OF IMPLEMENTATION PATH FOR FORMATIVE ASSESSMENT OF COLLEGE ENGLISH

A. IMPLEMENTATION PRINCIPLES

To ensure the scientificity, operability, and effectiveness of formative evaluation, the implementation process should follow the following six principles:

1. The principle of putting learning at the center: serving students' learning and development, rather than management or grading, with "promoting learning and teaching through evaluation" as the core value orientation.
2. Process closed-loop principle: Build a complete closed-loop of "goal implementation monitoring feedback improvement", avoid fragmented operations, and ensure the deep integration of evaluation and teaching.
3. Principle of Multiple Subjects: Integrate teacher evaluation, student self-evaluation, and peer evaluation to make students the main evaluators and cultivate reflective abilities.
4. Operational principle: Adapt to the actual teaching of large classes, reduce the burden on teachers, simplify the process through technical tools, and ensure that the evaluation can be implemented and promoted.
5. Motivational principle: Emphasize encouragement, pay attention to students' progress, protect learning confidence and enthusiasm, and avoid excessive judgment.
6. Principle of fairness and transparency: The evaluation criteria are open, the process is traceable, and the results are credible, avoiding subjective arbitrariness and ensuring fair and impartial evaluation.

B. OVERALL FRAMEWORK: "GOAL IMPLEMENTATION MONITORING FEEDBACK IMPROVEMENT" FIVE IN ONE

This article constructs a comprehensive implementation path for pre class diagnostic evaluation, in class process evaluation, and post class developmental evaluation. Formative evaluation accounts for 50% of the total semester grade, and final exams account for 50%. The implementation process and results are equally emphasized. The specific framework is as follows:

Objective layer: Clarify unit learning objectives and ability requirements, providing clear guidance for evaluation;

Implementation layer: Conduct evaluation and collect learning evidence throughout the entire process from pre class to in class to post class;

Monitoring layer: Real time tracking of learning progress to ensure standardized evaluation process and complete data;

Feedback layer: Timely, specific, and actionable feedback on evaluation results to guide improvement;

Improvement layer: Based on feedback, adjust teaching and learning strategies to form a continuous optimization loop.

C. PRE CLASS: DIAGNOSTIC EVALUATION - PRECISE POSITIONING, TEACHING BASED ON LEARNING

1) Objectives

Understand students' prior knowledge, skill levels, learning habits, interests, and difficulties, provide a basis for hierarchical teaching, goal setting, and task design, and avoid "one size fits all" teaching.

AI tools automatically review logs and reflections; teaching platforms archive portfolios and generate growth curves. In-class tasks are auto-checked to reduce teachers' workload by about 80%.

2) Implementation Content

1. Entrance/Unit Pre Test: Through vocabulary, grammar, listening, and reading basic level tests, accurately identify students' knowledge gaps and ability shortcomings;
2. Learning situation questionnaire: Investigate students' learning motivation, methods, difficulties, expectations, understand learning habits and personalized needs;
3. Pre study task inspection: Check students' completion of pre study and knowledge mastery through online pre study videos, word check-in, and pre study tests;
4. Establishment of learning records: Record students' starting level, provide a basis for comparison of subsequent progress, and form personalized learning records.

3) Evaluation Tools

Study Pass, Rain Classroom, Questionnaire Star, Unit Preview Sheet, Pre knowledge Test Paper, etc.

4) Application of Results

Adjust the key and difficult points and progress of teaching, and focus on explaining weak knowledge points; Implement hierarchical tasks and grouping to design differentiated learning tasks for students of different levels;

Early intervention for students with weak foundations, providing personalized tutoring and resource recommendations.

D. IN CLASS: PROCESS EVALUATION - REAL TIME TRACKING, INTERACTIVE ENHANCEMENT

In class is the core part of formative assessment, emphasizing immediate feedback, deep participation, and collabor-

orative interaction. Through collecting learning evidence in multiple ways, it achieves the goal of "promoting learning through evaluation".

1) Classroom Participation Evaluation

Classroom participation evaluation focuses on students' interaction quality and learning engagement, including:

Classroom answers, group discussions, presentations, role-playing, interactive quizzes, etc;

Using standardized evaluation scales, quantify from three dimensions: behavioral participation (number of speeches, frequency of interactions), cognitive participation (depth of thinking, logical expression), and emotional participation (cooperative attitude, positivity);

Pay attention to the quality of participation rather than just the frequency, encourage students to think actively, express themselves actively, and avoid "passive listening to lectures".

2) Evaluation of Oral Skills

Oral skills are the core weakness of college English teaching, and oral evaluation needs to be strengthened in class. Oral tasks such as impromptu answers, dialogue exercises, brief statements, and group debates;

Evaluate from four dimensions: fluency, accuracy, completeness, and interactivity, using a standardized scoring scale;

Encourage peer evaluation among students, allowing them to improve their oral expression skills through mutual correction.

3) Instant Task Evaluation

By reading, writing, and translating small tasks in class, achieve "completion in class, feedback in class, and correction in class":

The task design closely follows the unit objectives, with moderate difficulty, making it easy to complete in person;

Teachers demonstrate excellent answers, guide students to evaluate each other, and focus on explaining common errors;

Instant feedback on learning outcomes, helping students identify problems and correct errors in a timely manner.

4) Group Collaboration Evaluation

Group cooperation is an important way to cultivate students' ability to collaborate and communicate, and the evaluation dimensions include:

Task completion status, cooperative attitude, contribution level, language usage;

Combining self-evaluation within the group, peer evaluation between groups, and teacher evaluation to ensure comprehensive and objective evaluation;

Focus on students' teamwork, communication and expression, and sense of responsibility, rather than just task results.

5) Real Time Feedback Strategy

Multiple instant feedback methods are used in class:

Oral instant feedback: specific, positive, and actionable, such as "Your pronunciation is clear, but there are many grammar errors. It is recommended to pay more attention to tense consistency";

Board/screen feedback: Focus on explaining common errors to avoid repetitive problems;

Peer feedback: Pair checking, mutual correction, and enhancing students' reflective ability;

Tool feedback: Utilize Rain Classroom to instantly calculate the accuracy of answering questions and quickly grasp the learning situation of the class.

E. AFTER CLASS: DEVELOPMENTAL EVALUATION - CONSOLIDATION AND IMPROVEMENT, SELF REFLECTION

After class evaluation focuses on self-directed learning, continuous accumulation, reflection and improvement, promoting the extension of learning from the classroom to the outside, and cultivating students' ability for self-directed learning.

AI intelligent review tools are used to automatically extract keywords, identify common problems, generate class summary reports and personalized warning lists for learning logs and reflection notes. Teachers only need to provide targeted feedback to students with warnings. Smart teaching platforms automatically archive assignments, works, and test results, and generate growth curves and ability radar charts without manual sorting. In-class tasks and quizzes are automatically published, corrected, and fed back by the system. Under the traditional manual mode, teachers need to spend about 120 minutes per week; under the technology-enabled mode, only about 25 minutes per week is needed, reducing workload by about 80%. Reflection notes are simplified to weekly template-based entries; learning portfolios collect only key works; feedback adopts the mode of "common evaluation + sampling feedback + personalized push".

1) Homework Evaluation After Class

Homework evaluation not only looks at right and wrong, but also at attitude, progress, and correction:

Layered homework design: basic layer (consolidating knowledge), improvement layer (expanding abilities), and expansion layer (innovative applications), suitable for students of different levels;

Using comment based feedback to identify specific issues and provide improvement suggestions, avoiding simple right and wrong or scoring;

Pay attention to the quality of homework completion, correction status, and effort level, and encourage students to actively correct errors.

2) Self Directed Learning Evaluation

Self directed learning evaluation focuses on students' initiative and sustainability in extracurricular learning: Online

listening, reading, and word check-in tasks, recording learning duration, completion rate, and accuracy rate;

The evaluation dimensions include learning plan, execution intensity, persistence, and learning effectiveness; Utilize platform tools to automatically collect data and reduce the burden on teachers.

3) Learning Log/Reflection Notes

Guide students to record their learning content, gains, difficulties, and next steps on a weekly basis: Reflection notes require specific requirements, such as "This week I learned relative clauses and mastered the basic structure, but I still tend to confuse relative pronouns in writing. Next week, I will focus on practicing the application of relative clauses in writing;

Teachers regularly review and provide feedback, guiding students to actively identify problems, plan learning, and cultivate metacognitive abilities.

4) Project Based/Portfolio Evaluation

At the end of the semester, a project-based evaluation will be conducted to collect evidence of students' learning and growth

Theme projects: English speeches, short videos, book reports, short dramas, etc., showcasing language application abilities;

Learning portfolio: Collect the best works, reflection notes, and improvement drafts to visually demonstrate learning progress;

Focus on evaluating students' growth trajectory and level of effort, rather than absolute scores.

5) Periodic Self-Assessment and Peer Evaluation

Conduct a self-evaluation and peer evaluation for each unit:

Self evaluation: Students compare their learning goals and evaluate their achievement, level of effort, and weaknesses;

Mutual evaluation: evaluate each other's cooperation attitude, contribution, and progress within the group;

Help students become independent evaluators, enhance their reflective ability and learning initiative.

V. EMPIRICAL RESEARCH DESIGN

A. RESEARCH OBJECT

Select two natural classes for first-year non English majors at a certain university, with a total of 126 students.

Experimental group: 63 people, using the formative evaluation pathway constructed in this article;

Control group: 63 people, using traditional summative evaluation mode.

Informed consent was obtained from all participants. There was no significant difference ($p>0.05$) in English scores, gender, and major distribution between the two classes, indicating comparability.

To eliminate the testing effect, the control group received the same frequency of in-class exercises and quizzes as the experimental group, but without grading or feedback.

Parallel test papers with equivalent difficulty but different items were used for pre-test and post-test to avoid score improvement caused by question familiarity. A delayed post-test was conducted 8 weeks after the experiment to verify long-term ability retention.

B. RESEARCH VARIABLES

Independent variable: evaluation method (formative evaluation vs traditional evaluation);

Dependent variable:

1. College English academic performance (final unified exam score);

2. Learning motivation (intrinsic motivation, extrinsic motivation, learning engagement);

3. Self directed learning ability (planning, monitoring, regulating, and reflecting dimensions);

4. Classroom participation (behavioral participation, cognitive participation, emotional participation);

5. Language application ability (fluency, accuracy, and complexity in speaking and writing);

Control variables: textbooks, teachers, class hours, progress, and assessment content are consistent. Dependent variables include college English academic performance (final unified exam score), learning motivation, self-directed learning ability, classroom participation, and four core dimensions of language application ability: oral expression ability (fluency, accuracy, interactivity, pragmatic appropriateness), writing application ability (content completeness, logical coherence, language accuracy, format standardization), collaborative communication ability (communication, task participation, contribution), and self-directed learning ability (learning strategy, plan completion, reflection quality).

C. RESEARCH TOOLS

1. English academic performance: Entrance exam and final unified exam (out of 100 points);

2. Learning Motivation Scale: Adopting the adapted version of the English Learning Motivation Scale, it includes three dimensions: intrinsic motivation, extrinsic motivation, and learning engagement, with a total of 20 questions and a 1-5 point scoring system;

3. Self directed Learning Ability Scale: Using the "College English Self directed Learning Ability Scale", it includes 5 dimensions: planning, monitoring, regulation, evaluation, and reflection, with a total of 25 questions and a 1-5 point scoring system;

4. Classroom participation measurement scale: includes three dimensions: behavioral participation, cognitive participation, and emotional participation, with a total of 15 questions and a 1-5 point scoring system;

5. Language Application Ability Rating Scale: Spoken and written language are scored on a scale of 1-10, based on dimensions such as fluency, accuracy, completeness, and logical expression;

6. Interview outline: A semi-structured interview outline for teachers and students, used to obtain qualitative feedback.

All scales have good reliability (Cronbach's $\alpha > 0.82$) and validity, meeting statistical requirements. Objective verification indicators are added to form a dual evidence system of "subjective self-evaluation + objective performance". Objective data include platform learning duration, task completion rate, timely submission rate of assignments, error correction rate, and pre-class preview completion rate. Teachers conduct blind scoring on learning initiative, classroom participation, and task execution; group members conduct peer evaluation. The scale results are highly consistent with objective indicators (correlation coefficients > 0.70), proving that subjective self-evaluation can effectively reflect real ability changes.

D. RESEARCH HYPOTHESIS

1. Formative assessment can significantly improve students' academic performance in college English;

2. Formative evaluation can significantly enhance students' learning motivation;

3. Formative evaluation can significantly enhance students' ability for self-directed learning;

4. Formative evaluation can significantly improve students' classroom participation;

5. Formative evaluation can significantly enhance students' language application ability.

The research hypothesis is revised to: Formative assessment (process feedback + personalized guidance) improves learning behaviors, further enhances language application ability, and finally promotes academic performance. The causal chain eliminates circular reasoning.

VI. EMPIRICAL RESULTS AND ANALYSIS

A. COMPARISON OF ACADEMIC PERFORMANCE RESULTS

The Comparison of Academic Performance Results is shown in Table 1:

TABLE 1. Comparison of English Academic Scores Between Experimental Group and Control Group

Group	N	Pre-test Mean $\pm$ SD	Post-test Mean $\pm$ SD	t	p
Experimental Group	63	71.2 $\pm$ 6.8	78.6 $\pm$ 5.9	6.72	< 0.001
Control Group	63	70.8 $\pm$ 7.1	72.5 $\pm$ 6.6	1.35	> 0.05

The results indicate that:

There was no significant difference between the pre-test and control groups in the experimental group ($p > 0.05$), but the post test was significantly higher than the control group ($p < 0.001$);

The experimental group showed an average improvement of 7.4 points, with an increase of approximately 10.4%; The control group only improved by 1.7 points, with an increase of about 2.4%.

Assumption 1 is true: Formative evaluation significantly improves academic performance, and its process oriented approach effectively avoids students' "pre exam rush" and promotes continuous learning and knowledge accumulation.

B. COMPARISON OF LEARNING MOTIVATION RESULTS

The Comparison of Learning Motivation Results is shown in Table 2:

TABLE 2. Comparison of Learning Motivation Between Experimental Group and Control Group

Group	N	Pre-test Mean $\pm$ SD	Post-test Mean $\pm$ SD	t	p
Experimental Group	63	3.12 $\pm$ 0.41	3.86 $\pm$ 0.35	7.11	< 0.001
Control Group	63	3.08 $\pm$ 0.43	3.21 $\pm$ 0.39	0.89	> 0.05

The results indicate that:

The learning motivation of the experimental group significantly improved, with a significant increase in intrinsic interest, learning engagement, and persistence;

There was no significant change in the control group, and external motivation (exams, credits) still dominated.

Assumption 2 is true: Formative evaluation effectively stimulates students' intrinsic learning motivation and alleviates the phenomenon of "passive learning" through continuous feedback and positive incentives.

C. COMPARISON OF SELF-DIRECTED LEARNING ABILITY RESULTS

The comparison of Self Directed Learning Ability results is shown in Table 3:

TABLE 3. Comparison of Autonomous Learning Ability Between Experimental Group and Control Group

Group	N	Pre-test Mean $\pm$ SD	Post-test Mean $\pm$ SD	t	p
Experimental Group	63	3.05 $\pm$ 0.38	3.79 $\pm$ 0.33	6.95	< 0.001
Control Group	63	3.01 $\pm$ 0.40	3.15 $\pm$ 0.37	1.02	> 0.05

The results indicate that:

The experimental group has significantly improved in terms of planning, monitoring, regulation, and reflection, and students are more proactive in arranging learning, identifying problems, and improving methods; The control group's autonomous learning ability did not show significant changes and still relied on teacher arrangements.

Assumption 3 is true: Formative evaluation effectively cultivates students' ability to learn independently through learning logs, self check-ins, self-evaluation and reflection, and helps students transition from "passive acceptance" to "active learning".

D. COMPARISON OF CLASSROOM PARTICIPATION RESULTS

The Comparison of Classroom Participation Results is shown in Table 4:

TABLE 4. Comparison of Classroom Participation Between Experimental Group and Control Group

Group	N	Pre-test Mean ± SD	Post-test Mean ± SD	t	p
Experimental Group	63	3.06 ± 0.42	3.81 ± 0.36	7.32	< 0.001
Control Group	63	3.03 ± 0.41	3.18 ± 0.38	1.15	> 0.05

The results indicate that:
The experimental group students are more willing to take the initiative to speak, participate in group discussions, and present reports, and their behavioral, cognitive, and emotional participation have significantly improved;
The classroom participation of the control group did not show significant changes, and still mainly relied on passive listening, with a relatively low proportion of active participation.
Assumption 4 is true: Formative evaluation incorporates classroom participation into performance assessment, effectively mobilizes students' classroom participation enthusiasm through immediate feedback and positive incentives, and creates an active classroom atmosphere.

VII. CONCLUSION

The core mechanism by which formative evaluation can significantly improve the effectiveness of college English teaching based on comprehensive quantitative and qualitative results is:

1. Process oriented approach to solving the dilemma of exam taking: Formative evaluation permeates the entire teaching process, mitigating the reliance on intensive exam-oriented study while facilitating the continuous accumulation of technical knowledge and the gradual refinement of professional skills required for the industrial sector. This persistent assessment model ensures students develop the rigorous analytical capabilities necessary to address the evolving complexities of fiber research and modern industrial applications;
2. Multiple feedback strengthens the motivation for improvement: Timely, specific, and actionable feedback allows students to clarify "where they went wrong and how to improve", forming a virtuous cycle of "evaluation feedback improvement";
3. Subject participation stimulates intrinsic motivation: self-evaluation, peer evaluation, project presentation, and other activities transform students from "passive evaluation objects" to "active evaluation subjects", stimulating their intrinsic learning interest and initiative;
4. Ability oriented improvement of evaluation content: Highlighting the evaluation of core abilities such as speaking, writing, and collaboration, promoting the transformation of teaching from "knowledge imparting" to "ability cultivation", and enhancing language application ability.

At the same time, the interview results also showed that there are still some details in the implementation of formative evaluation: some students do not attach enough importance to self-evaluation and peer evaluation, and the objectivity of feedback needs to be improved; Under large

class teaching, the coverage of personalized feedback still needs to be expanded; The feedback and evaluation design abilities of some teachers need further improvement. These issues will be addressed in the subsequent optimization strategies.

AUTHOR CONTRIBUTIONS

Conceptualization –Yuqing Zhang and Zhiyuan Jin; methodology – Yuqing Zhang and Zhiyuan Jin; investigation – Yuqing Zhang and Zhiyuan Jin; writing-original draft preparation – Yuqing Zhang and Zhiyuan Jin. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research was supported by the Postgraduate Research Fund Project of Xi'an International Studies University in 2025, China, under Grant No. 2025ZC003: Research on the Construction and Implementation of a Human-Computer Collaboration Model for College English Teaching Empowered by Large Language Models.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] Y. Deng and Y. Liu, "The intervention effect of integrating mental health nursing knowledge into college english teaching on students with emotional disorders," *Schizophrenia Bulletin*, vol. 52, no. Supplement_1, pp. S86-S87, 2026, doi: 10.1093/SCHBUL/SBAG003.126.

[2] X. Wang, M. Liu, and P. Yan, "Countermeasures for Improving College English Teaching Efficacy in Private Universities Based on Big Data," *Education Reform and Development*, vol. 8, no. 1, pp. 180-185, 2026, doi: 10.26689/ERD.V8I1.13798.

[3] Q. Song, "Research on Strategies for Integrating Traditional Chinese Culture into College English Teaching," *Journal of Contemporary Educational Research*, vol. 10, no. 1, pp. 133-138, 2026, doi: 10.26689/JCER.V10I1.13693.

[4] L. Wei, D. Hui, and Y. Yin, "Intelligence-Driven Multi-Dimensional Collaborative Model for Blended University English Teaching: Design, Implementation, and Effectiveness Evaluation Based on Digital Web Technologies," *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT)*, vol. 21, no. 1, pp. 1-17, 2026, doi: 10.4018/IJWLTT.400754.

[5] X. Huang, "Optimizing the Integration Pathways of Humanistic Education in College English Teaching Under the Guidance of Core Literacy," *Journal of Modern Education and Culture*, vol. 2, no. 11, 2026, doi: 10.70767/JMEC.V2I11.866.

[6] W. Jie and M. Xiangni, "Computational Methods for Intuitionistic Fuzzy Group Decision-Making: Evaluating College English Teaching Quality," *International Journal of Fuzzy System Applications (IJFSA)*, vol. 15, no. 1, pp. 1-18, 2026, doi: 10.4018/IJFSA.399757.

[7] Y. Duan, "Pathways for Digital Empowerment to Enhance the Quality of College English Teaching," *Exploration of Educational Management*, vol. 4, no. 1, 2026, doi: 10.12417/3029-2328.26.01.007.

[8] T. Jraiban A F, Y. Sawalh M, R. Shahwan A A, et al., "Requirements for Employing Artificial Intelligence Applications in English Language Education at Universities: Perspectives of Faculty Members," *Theory and Practice in Language Studies*, vol. 16, no. 1, pp. 292-303, 2026, doi: 10.17507/TPLS.1601.32.

[9] X. Che, "Research on Integrating Outstanding Traditional Chinese Culture into College English Teaching," *Journal of Contemporary Educational Research*, vol. 9, no. 12, pp. 292-296, 2025, doi: 10.26689/JCER.V9I12.13311.

- [10] Y. Lin, "Exploration of the Cultivation Mode of Critical Thinking Ability in College English Teaching Under the Background of the Belt and Road," *Education Reform and Development*, vol. 7, no. 12, pp. 220-225, 2025, doi: 10.26689/ERD.V7I12.13533.
- [11] J. Zhang, "Research on College English Teaching and Application Based on the ADDIE Model from the Perspective of Technological Empowerment," *Journal of Higher Education Research*, vol. 6, no. 6, 2025, doi: 10.32629/JHER.V6I6.4699.
- [12] Y. Wang, "Research on the Practical Mechanism of Integrating Curriculum Ideology and Politics into College English Teaching," *Journal of Sociology and Education*, vol. 1, no. 12, 2025, doi: 10.63887/JSE.2025.1.12.27.
- [13] Y. Zhou, Y. Wang, and Y. Zhang, "Exploration of AI-Enabled Ideological and Political Education with Military Cases in College English Teaching," *Journal of Education and Educational Research*, vol. 16, no. 3, pp. 90-96, 2025, doi: 10.54097/01R23K28.
- [14] H. Wei, "Research on The Cultivation and Dissemination of Chinese Cultural Awareness in College English Teaching," *Education Reform and Development*, vol. 7, no. 11, pp. 79-84, 2025, doi: 10.26689/ERD.V7I11.13059.
- [15] C. Tan, "Empowerment and Reconstruction: A Study on the Application Models and Effectiveness of AI in College English Teaching for Science Students," *International Journal of Social Science and Education Research*, vol. 8, no. 12, pp. 340-345, 2025, doi: 10.6918/IJOSSER.202512_8(12).0048.
- [16] X. Zhang, H. Tang, H. Zhang, et al., "Utilizing Multiple Intelligences Theory in College English Teaching: Strategies to Enhance Low-Proficiency Learners' English Language Skills," *Journal of English Language Teaching and Applied Linguistics*, vol. 7, no. 8, pp. 52-63, 2025, doi: 10.32996/JELTAL.2025.7.8.6.
- [17] M. Xiaomei, H. Jing, and L. Zhaoyi, "A Study of "Chinese Cultural Aphasia" in College English Teaching from a Intercultural Perspective," *Journal of International Social Science*, vol. 2, no. 12, pp. 177-186, 2025, doi: 10.62639/SSPJISS28.20250212.
- [18] Y. Song, "Exploration of the Narrative Ability Cultivation Model for Telling Chinese Stories in English in College English Teaching," *Journal of Social Science and Cultural Development*, vol. 2, no. 7, 2025, doi: 10.70767/JSSCD.V2I7.776.
- [19] L. Bu and L. Song, "Analysis on Methods and Effect of Cultural Introduction in College English Teaching," *Journal of Higher Education Teaching*, vol. 2, no. 6, 2025, doi: 10.62517/JHET.202515614.
- [20] X. Zhao, "Research on the Integration of Local Intangible Cultural Heritage into College English Teaching Based on Output Orientation," *New Explorations in Education and Teaching*, vol. 3, no. 11, 2025, doi: 10.70711/NEET.V3I11.8113.