An Empirical Study on the Improvement of Higher Vocational College Students' Workplace English Writing Ability by Deep Learning-Driven Intelligent Marking System

Hui Li, Yongmei Zhou
School of Fundamental Education, Beijing Polytechnic College, Beijing 100042, China
Corresponding author: Hui Li (e-mail: amier_Lee@163.com)

ABSTRACT To address low grading efficiency, delayed feedback, and insufficient personalized guidance in higher vocational workplace English writing, this study constructs a deep learning-driven intelligent marking system and verifies its teaching effect through empirical research. A total of 120 students from two higher vocational English classes were divided into an experimental group using the intelligent grading system and a control group using conventional teacher grading. During the 16-week experiment, writing tests, questionnaires, interviews, and learning records were collected and statistically analyzed. The system integrates BERT, LSTM, natural language processing, and workplace-document evaluation rules to assess grammatical accuracy, sentence diversity, content relevance, and workplace standardization. The results show that the experimental group significantly outperformed the control group in overall writing scores and all core dimensions, while writing interest, autonomous learning, and confidence also improved. The proposed system reduces teacher workload and provides a data-driven feedback pipeline suitable for smart educational environments based on networked human-machine interaction.

INDEX TERMS deep learning,intelligent grading,workplace English writing,natural language processing,human-machine interaction

I. INTRODUCTION

A. RESEARCH BACKGROUND

TO implement the National Vocational Education Reform Implementation Plan and further improve the quality of talent training, the Ministry of Education promulgated the English Curriculum Standard for Higher Vocational Education (Junior College) (2021 Edition) on March 23, 2021, clearly defining the teaching objectives of higher vocational public English courses [1]. It aims to enhance students' crosscultural communication ability in the workplace, promote effective communication in a multicultural context, improve language thinking ability, and perfect autonomous learning ability, so as to comprehensively cultivate students' core literacy in these key fields [2-4]. Workplace English writing is a core module of higher vocational English teaching, which is directly related to students' professional competitiveness and employment quality [5,6].

At present, the improvement of higher vocational students' Workplace English writing ability faces many challenges, such as disconnection between teaching content and workplace needs, outdated teaching methods, insufficient students' learning motivation, and lack of teaching resources [7]. With the maturity of artificial intelligence technologies such as natural language processing in China, some intelligent grading software has emerged, providing important reference for composition grading [8,9]. As an independent grading subject, intelligent grading software brings new thinking to composition grading [10]. The state fully promotes the deep integration of information technology with education and teaching, encourages schools to use big data technology to collect, analyze and feedback education and teaching activities and students' behavioral data, and provides support for personalized learning and targeted teaching [11,12]. It supports all kinds of schools at all levels to build smart campuses, and comprehensively uses Internet, big data, artificial intelligence and virtual reality technologies to explore future education and teaching models [13].

Most Workplace English writing compositions used to be graded manually, which is a heavy and time-consuming task. Especially when there are a large number of students and a

huge grading workload, manual grading by teachers is inefficient, and the grading results cannot be fed back to students in a timely manner. With the passage of time, learners may forget the details of their assignments. In addition, due to the subjectivity of graders and the non-uniqueness of answers, misjudgment, over-scoring and under-scoring may occur. If a deep learning-based intelligent grading system can be used to automatically complete these complex tasks, the efficiency of assignment grading will be greatly improved, learners can get timely evaluation feedback, and the quality of scoring will not be affected by human factors [14,15].

At present, some studies have explored the application of intelligent grading systems in English teaching, but empirical studies focusing on higher vocational students' workplace writing ability are relatively scarce. Therefore, exploring the improvement effect of the deep learning-driven intelligent grading system on higher vocational students' Workplace English writing ability has important practical significance and practical value.

B. RESEARCH STATUS

Artificial intelligence is a challenging interdisciplinary science. Artificial intelligence researchers need to integrate knowledge of computer science, mathematics, information science, cognitive science and other disciplines [16]. Generally speaking, the main goal of artificial intelligence research is to understand the essence of intelligence, and design computer application programs artificially to enable machines to handle complex tasks that can only be accomplished by human intelligence. The goal of artificial intelligence is divided into two parts: "artificial" and "intelligence". "Artificial" means that computer applications replace human subjects, while the complex tasks referred to by "intelligence" are issues that researchers need to dynamically think about combined with the era and technological production level. For example, modern computers can already perform and achieve better results in heavy engineering calculations and reasoning decisions that needed to be done manually in the past. Therefore, researchers need to reconsider those "complex tasks that require human intelligence to complete", and constantly change the direction and goals of artificial intelligence research combined with the development of the times and technological progress.

Deep learning (also known as deep structure learning or hierarchical learning) originates from the research of artificial neural networks in the field of machine learning. It is a general algorithm based on learning data representation, rather than an algorithm for specific tasks. Deep learning is a branch of machine learning development [17]. Deep learning uses multiple processors with complex structures or composed of multi-layer nonlinear transformations to combine low-level intuitive features into sparse high-level abstract features, and represents original data with distributed feature attributes.

With the advent of the intelligent era, "Automated Essay Evaluation (AEE)" or "Automated Writing Evaluation (AWE; Automated Essay Scoring (AES))" has gradually developed into intelligent composition grading [18]. It is a way of systematically applying intelligent technology to grade compositions to obtain scores, comments and other feedback content, helping teachers improve grading efficiency, and further enhancing students' writing passion and ability. The development and application of intelligent grading abroad can be traced back to the 1960s, and intelligent grading systems such as E-Rater, IntelliMetric and PEG have been applied in large-scale high-stakes tests [19].

In the field of basic education in China, the application of intelligent grading in English subjects is the earliest and the most extensive [20]. This is because China has the earliest involvement in intelligent grading, with abundant developed systems and applied theoretical research results. In view of the value of intelligent composition grading in teaching, some scholars have explored the application mode of intelligent grading. Sui Xiaobing and Cheng Lulu proposed an interactive teaching mode of college English writing MOOC assisted by intelligent grading, including "MOOC learning, extensive reading", "brainstorming, flipped classroom", "online collaboration, repeated revision", "peer evaluation, teacher comment", "in-depth revision, teacher summary" [21]. Similarly, based on previous research results and teaching experience, Tang Jinlan proposed a pilot mode of applying intelligent grading to teaching, and verified and optimized the mode through controlled experiments [22].

C. RESEARCH IDEAS

(1) Sort out relevant theories and research status, and clarify the core functions and application advantages of the deep learning-driven intelligent grading system;

(2) Construct a deep learning intelligent grading system suitable for higher vocational Workplace English writing, and clarify the system architecture, model and grading dimensions;

(3) Select higher vocational students as research subjects, carry out a 16-week empirical experiment, divide them into experimental group and control group, and adopt intelligent grading system-assisted teaching and traditional teacher grading teaching respectively;

(4) Use statistical analysis methods to explore the improvement effect of the intelligent grading system on higher vocational students' Workplace English writing ability, and summarize the advantages and disadvantages in system application.

D. RESEARCH INNOVATIONS

First, focus on the subdivision field of higher vocational Workplace English writing, break through the limitation that existing studies mostly pay attention to general English writing, combine the "workplace-oriented" characteristics of higher vocational education, construct a deep learning intelligent grading system suitable for higher vocational students, highlight the grading dimensions such as format

specifications, industry terms and cross-cultural adaptation of workplace documents, and fit the actual post needs;

Second, adopt BERT model, LSTM model and other models to construct an intelligent grading model, improve the accuracy and pertinence of grading, and make up for the deficiencies of traditional intelligent grading systems;

Third, carry out empirical research to comprehensively explore the impact of the intelligent grading system on students' writing ability, learning interest and autonomous learning ability;

Fourth, pay attention to teachers' teaching experience, which can provide practical reference for the teaching reform of Workplace English writing in higher vocational colleges.

II. CONSTRUCTION OF DEEP LEARNING-DRIVEN INTELLIGENT GRADING SYSTEM

A. SYSTEM CONSTRUCTION OBJECTIVES

The core objectives of the deep learning-driven intelligent grading system constructed in this study include three aspects:First, improve grading efficiency, realize automatic grading of higher vocational Workplace English writing texts, reduce teachers' grading workload, and ensure timely feedback;Second, improve grading quality, realize multi-dimensional accurate grading of writing texts, which can not only identify basic errors, but also evaluate the semantics, logic and workplace standardization of texts;Third, promote the improvement of students' abilities, help students independently find problems and improve deficiencies through instant feedback, personalized practice, data tracking and other functions, improve Workplace English writing ability and autonomous learning ability, and cultivate students' workplace literacy at the same time, fitting the talent training objectives of higher vocational colleges.

To ensure practical utility, the system was designed with an intuitive interface focusing on low cognitive load and high accessibility for users with varying levels of digital literacy.

B. OVERALL SYSTEM ARCHITECTURE

The deep learning-driven intelligent grading system constructed in this study is divided into four layers: data, technology, application, and user layers. These layers operate through a synchronous data-exchange protocol: when a user selects a specific document type (e.g., a business email) in the Application Layer, this metadata is transmitted to the Technology Layer. This trigger automatically invokes the dynamic threshold matching strategy, which pulls the corresponding industry-specific grading parameters from the Data Layer to calibrate the BERT model's scoring sensitivity in real-time.

C. DATA LAYER

The data layer is the foundation of the system, which is mainly responsible for collecting, storing and managing various data required for system operation:

- Corpus data is the core data, including corpus related to higher vocational Workplace English writing and error corpus, which are used for pre-training and optimization of deep learning models to ensure the accuracy and pertinence of grading;
- Student data includes students' basic information, writing tasks, grading results, error records, learning duration, etc., which are used for personalized feedback and data tracking;
- Teacher data includes teachers' basic information, teaching tasks, grading records, teaching data reports, etc., which are used for teachers to carry out targeted teaching and teaching evaluation;
- System data includes system operation logs, parameter settings, etc., which are used for system maintenance and optimization.

The data layer utilizes distributed storage technology to ensure high availability and data redundancy. To strictly protect students' personal data and academic achievements, the system implements Advanced Encryption Standard (AES-256) for data-at-rest and robust anonymization protocols during the pre-processing phase to ensure privacy compliance.

D. TECHNOLOGY LAYER

The technology layer is the core of the system, which is mainly responsible for providing core technical support required for system operation:

- Deep learning models include BERT model and LSTM model, which are used for text scoring and error recognition respectively;
- Natural language processing technology is used for text word segmentation, part-of-speech tagging, syntactic analysis, semantic understanding, etc., to achieve accurate analysis of writing texts;
- Big data analysis technology is used to analyze students' writing data and teachers' teaching data, providing data support for personalized feedback and teaching guidance;
- Network security technology is used to ensure the operation safety of the system, prevent data leakage, malicious attacks and other problems, and ensure the stable operation of the system.

The technology layer adopts modular design, which is convenient for subsequent system upgrade and function optimization, and integrates a dynamic threshold matching strategy to enhance the applicability of the system.

E. APPLICATION LAYER

The application layer is the function realization layer of the system, which is mainly responsible for realizing core functions such as intelligent grading, feedback guidance, data management and teaching assistance:

(1) Writing submission module: Students can submit workplace English writing tasks through web and mobile

terminals, supporting multiple submission methods such as text input and file upload, adapting to different learning scenarios; At the same time, the system supports the selection of multiple workplace document types. Students can select the corresponding type according to teaching task requirements, and the system will grade according to the specifications of this type.

(2) Intelligent grading module: After receiving the writing text submitted by students, the system conducts multi-dimensional grading of the text through natural language processing technology and deep learning models; At the same time, it evaluates the grammatical accuracy, sentence diversity, content relevance and workplace standardization of the text, generates a comprehensive score and a detailed error analysis report, and the grading time does not exceed 30 seconds to ensure timely feedback.

(3) Personalized feedback module: According to the intelligent grading results and students' historical writing data, the system generates personalized feedback suggestions for students, including error cause analysis, modification suggestions, targeted exercises, etc.; At the same time, the system pushes workplace writing model essays and knowledge points related to students' weak links to help students consolidate and improve.

(4) Data management module: Responsible for managing various data of students, teachers and the system, including students' writing data, teachers' grading data, system operation data, etc.; At the same time, the system automatically generates student learning reports and teacher teaching reports, intuitively showing students' learning progress and teaching effect.

(5) Teaching assistance module: Provide teaching support for teachers, including functions such as teaching task release, assignment, grading management, targeted guidance, etc. Teachers can view students' writing data and grading results through the system; At the same time, the system provides workplace writing teaching resources to help teachers optimize teaching plans and improve teaching quality.

F. USER LAYER

The user layer is the operation entrance of the system, mainly facing two types of users: higher vocational students and English teachers:

- Student users can submit writing tasks, view grading results, obtain personalized feedback, conduct targeted exercises, etc. through the user layer;
- Teacher users can release teaching tasks, view students' writing data, manage grading tasks, generate teaching reports, etc. through the user layer.

The system interface is simple in design and convenient in operation, and supports multi-terminal access, facilitating users to carry out learning and teaching activities anytime and anywhere.

G. REALIZATION OF CORE SYSTEM FUNCTIONS

1) Realization of Intelligent Grading Function

The intelligent grading function is the core function of the system, mainly realized through the collaborative work of BERT model and LSTM model. The models were pre-trained on a hybrid dataset comprising the WikiText-103 corpus and a proprietary collection of 50,000 annotated higher vocational English essays. For the BERT model, we utilized a learning rate of 2×10^{-5}, a batch size of 32, and trained for 5 epochs. The LSTM model utilized 256 hidden units with a dropout rate of 0.2. Validation tests showed that the error recognition accuracy reached 91.2% and the scoring correlation (QWK) with human experts reached 0.88. The specific implementation process is as follows:

First, after students submit writing texts, the system preprocesses the texts through natural language processing technology and converts the texts into a format recognizable by the model;

Second, the LSTM model performs error recognition on the preprocessed text, captures grammatical errors, spelling errors, sentence errors, etc. in the text, and generates an error list and error cause analysis; Then, the BERT model performs semantic understanding and multi-dimensional scoring on the text, combined with the format specifications and industry requirements of workplace documents, scoring from four dimensions: grammatical accuracy, sentence diversity, content relevance, and workplace standardization, each accounting for 25%, and the comprehensive score is the student's writing score;

Finally, the system integrates error analysis and scoring results to generate a detailed grading report and feed it back to students and teachers.

In addition, the system introduces a dynamic threshold matching strategy to dynamically adjust the scoring threshold according to the difficulty and type of writing tasks to ensure the fairness and accuracy of scoring. The scoring threshold for basic workplace documents (such as simple business emails) is set to 0.75, and the scoring threshold for complex workplace documents (such as foreign trade contracts) is set to 0.85.

2) Realization of Personalized Feedback Function

The personalized feedback function is based on students' writing data and grading results, and uses big data analysis technology to mine students' weak links and generate targeted feedback suggestions. The specific implementation process is as follows:

First, the system collects students' historical writing data, including error types, score changes, learning duration, etc. of each writing;

Second, analyze students' weak links through big data analysis technology;

Then, combined with students' weak links, the system generates a personalized error analysis report, detailing error causes, modification methods, and pushes relevant knowledge points and exercise questions; Finally, students

can conduct targeted exercises according to feedback suggestions, and the system tracks the exercise effect in real time, adjusts feedback content, and forms a closed loop of "grading-feedback-practice-improvement".

3) Realization of Data Management and Analysis Function

The data management and analysis function is mainly responsible for collecting, sorting and analyzing various data during system operation, providing data support for students' learning and teachers' teaching. The specific implementation is as follows:

First, the system adopts distributed storage technology to store various data of students, teachers and the system to ensure data security and integrity;

Second, through big data analysis technology, statistically analyze students' writing data, including error type distribution, score change trend, learning duration statistics, etc., generate student learning reports, and intuitively show students' learning progress;

Then, statistically analyze teachers' teaching data, including grading workload, students' average scores, common error distribution, etc., generate teacher teaching reports to help teachers fully understand teaching effect and optimize teaching plans.

At the same time, the system supports data query, export and other functions. Teachers and students can query relevant data and export reports according to needs, providing convenience for learning and teaching.

III. EMPIRICAL RESEARCH DESIGN AND IMPLEMENTATION

A. RESEARCH HYPOTHESES

Based on relevant theories and research background, this study puts forward the following three research hypotheses:

H1: The use of deep learning-driven intelligent grading system to assist teaching can significantly improve higher vocational students' comprehensive scores in Workplace English writing.

H2: The use of deep learning-driven intelligent grading system to assist teaching can significantly improve the scores of the four core dimensions of higher vocational students' Workplace English writing ability.

H3: The deep learning-driven intelligent grading system can improve higher vocational students' writing interest, autonomous learning ability and workplace writing confidence.

B. RESEARCH SUBJECTS

This study selected a total of 120 students from two classes of Grade 2024 higher vocational business administration major in a higher vocational college as research subjects, among which Class 1 (60 students) was the experimental group and Class 2 (60 students) was the control group.

The selection basis is as follows:

First, the two classes are both higher vocational business administration major, with basically the same professional background and English foundation, both offering the course Higher Vocational Workplace English writing, with the same teaching progress, teaching content and teachers, so they are comparable;

Second, before the experiment, the Workplace English writing ability of the two classes was tested. Through statistical analysis, there was no significant difference in the comprehensive writing scores and scores of each dimension between the two classes ($P > 0.05$), ensuring the fairness of the experiment. This study was approved by the Ethics Committee of Beijing Polytechnic College, and was performed in accordance with the principles of the Declaration of Helsinki. All eligible participants signed an informed consent form. The basic information of the research subjects is as follows:

- In the experimental group of 60 students, there were 22 males and 38 females; aged 18-20 years old, with an average age of 19.2 years old; medium English foundation, including 28 students with English Level 3 and above, and 32 students below Level 3.
- In the control group of 60 students, there were 20 males and 40 females; aged 18-20 years old, with an average age of 19.1 years old; medium English foundation, including 26 students with English Level 3 and above, and 34 students below Level 3.

There was no significant difference in the basic information of the two groups of students ($P > 0.05$), which was suitable for empirical research.

C. RESEARCH VARIABLES

1) Independent Variable

The independent variable of this study is the teaching method of Workplace English writing, divided into two types:

- The experimental group adopts intelligent grading system-assisted teaching. After completing writing tasks, students submit them to the system, and the system automatically grades and generates feedback;
- The control group adopts the traditional teaching mode. After completing writing tasks, students are manually graded by teachers, and then centralized commentary and individual guidance are carried out.

The teaching content, teaching progress, and teachers of the two teaching methods remain consistent. To control for confounding variables, both groups were strictly limited to a maximum of three revision cycles per writing task, ensuring that the observed differences were due to the source of feedback (intelligent system vs. teacher) rather than the frequency of revision.

2) Dependent Variable

The dependent variable of this study is higher vocational students' Workplace English writing ability, specifically including two aspects:

First, comprehensive writing scores, obtained through writing tests before and after the experiment; Second, scores

of the four core dimensions of writing ability (grammatical accuracy, sentence diversity, content relevance, workplace standardization), obtained by scoring the writing test papers.

In addition, the dependent variables also include students' writing interest, autonomous learning ability and workplace writing confidence, obtained through questionnaires and interviews.

3) Control Variables

To ensure the scientificity and reliability of the experimental results, this study controls the following variables:

First, teaching variables: the teaching content, teaching progress, teachers and teaching duration of the two classes remain consistent, both carry out teaching in accordance with the syllabus of Higher Vocational Workplace English writing, focusing on explaining writing skills of workplace documents such as business emails, resumes, meeting minutes and foreign trade correspondence;

Second, student variables: during the experiment, students in both classes do not participate in other English writing training or tutoring to ensure the consistency of the learning environment;

Third, test variables: the same set of test questions is used for writing tests before and after the experiment, with unified scoring standards. Two senior higher vocational English teachers score independently according to the scoring rules. The inter-rater reliability was measured using the Pearson correlation coefficient, which reached 0.92 for both pre-test and post-test assessments. Furthermore, a double-blind scoring procedure was implemented where the instructors were blinded to the group assignments of the students to prevent any potential evaluation bias.

D. RESEARCH TOOLS

1) Writing Test Papers

This study designed two sets of writing test papers for pre-test and post-test. The difficulty, question types and scoring standards of the two sets of papers are consistent, both focusing on common document types in higher vocational Workplace English writing (such as business emails, resumes, etc.).

The scoring standard is divided into four dimensions: grammatical accuracy (25 points), sentence diversity (25 points), content relevance (25 points), workplace standardization (25 points), with a total score of 100 points.

2) Questionnaire

After the experiment, a Questionnaire on Higher Vocational Students' Workplace English writing Learning was designed to conduct a questionnaire survey on the experimental group students, so as to understand the changes in students' experience with the deep learning-driven intelligent grading system, learning interest, autonomous learning ability and other aspects.

The questionnaire contains a total of 20 questions, adopting the Likert 5-point scale (1=strongly disagree, 2=disagree, 3=neutral, 4=agree, 5=strongly agree), divided into four dimensions: system use experience (5 questions), writing interest (5 questions), autonomous learning ability (5 questions), workplace writing confidence (5 questions).

After testing, the Cronbach's α coefficient of the questionnaire is 0.86, indicating that the questionnaire has good reliability; through expert validity test, the content validity index of the questionnaire is 0.91, indicating that the questionnaire has good validity and can effectively collect relevant data.

E. EXPERIMENTAL IMPLEMENTATION PROCESS

The experiment lasted for 16 weeks, divided into three parts: experimental preparation stage, experimental implementation stage and experimental summary stage. The specific implementation process is as follows:

1) Experimental Preparation Stage (Week 1)

(1) Complete the debugging and optimization of the deep learning-driven intelligent grading system to ensure the normal operation of the system;

(2) Conduct pre-test on students of the two classes, requiring students to complete workplace English writing tasks within the specified time;

(3) Conduct training on the use of the intelligent grading system for students in the experimental group, explaining the operation method, function modules, precautions, etc. of the system to ensure that each student can use the system proficiently;

(4) Formulate an experimental teaching plan, clarify the teaching content, teaching progress, teaching methods, etc. of the two classes to ensure the consistency of the teaching process.

2) Experimental Implementation Stage (Weeks 2-15)

The experimental group adopts deep learning-driven intelligent grading system-assisted teaching, and the specific teaching process is as follows:

Initially, teachers distribute workplace writing assignments with explicit requirements and scoring rubrics. Subsequently, students perform independent writing tasks and upload them to the intelligent platform. Upon submission, the system executes automated grading to provide instantaneous multi-dimensional feedback. Furthermore, students leverage these diagnostic reports to iteratively refine their work. Concurrently, instructors monitor real-time data trends to provide targeted interventions for individual weaknesses during weekly personalized practice sessions.

The control group adopts the traditional teaching mode, and the specific teaching process is as follows:

(1) Teachers release workplace writing tasks, clarifying writing requirements, scenarios and scoring standards;

(2) Students independently complete writing tasks and submit them to teachers;

(3) Teachers manually grade students' compositions, mark errors, give scores and modification suggestions;

(4) Teachers give centralized commentary on students' compositions in class, focusing on explaining common errors, and give individual guidance to individual students;

(5) Students revise their compositions according to teachers' grading suggestions and submit them to teachers for re-grading until they meet the requirements;

(6) Arrange 1 centralized practice per week, teachers assign exercise questions, and after students complete them, teachers conduct centralized grading and commentary.

3) Experimental Summary Stage (Week 16)

(1) Conduct post-test on students of the two classes, using the same writing test paper as the pre-test, requiring students to complete workplace English writing tasks within the specified time;

(2) Distribute questionnaires to students in the experimental group, a total of 60 questionnaires were distributed, and 58 valid questionnaires were recovered;

(3) Collect and sort out various data in the experimental process, including writing test scores, questionnaire data, interview records, students' writing data, teachers' teaching data, etc.

IV. EMPIRICAL RESULTS AND ANALYSIS

Through a 16-week empirical experiment, this study obtained the following conclusions, with specific analysis as follows.

A. STATISTICAL ANALYSIS OF WRITING SCORES BEFORE AND AFTER THE EXPERIMENT

The comprehensive writing scores of the experimental group and the control group before and after the experiment were analyzed, and the results are shown in Table 1 below.

It can be seen from Table 1 that before the experiment, the comprehensive Workplace English writing ability of the two groups of students was at the same level with no significant difference, showing good comparability. After the experiment, the comprehensive writing scores of both groups improved, but there was a significant difference in the improvement range.

- The average post-test score of the experimental group was 78.62 points, an increase of 16.27 points compared with the pre-test, with an increase rate of 26.1%;
- The average post-test score of the control group was 68.95 points, an increase of 7.07 points compared with the pre-test, with an increase rate of only 11.4%.

The independent sample t-test results show that there is an extremely significant difference in the post-test scores between the two groups (t=9.25, $P = 0.000 < 0.01$), indicating that the improvement effect of the experimental group's comprehensive writing scores is far better than that of the control group.

This result verifies research hypothesis H1, that is, the use of deep learning-driven intelligent grading system to assist teaching can significantly improve higher vocational students' comprehensive scores in Workplace English writing.

B. STATISTICAL ANALYSIS OF SCORES IN EACH DIMENSION OF WRITING ABILITY

The scores of the four core dimensions of writing ability of the experimental group and the control group before and after the experiment, including grammatical accuracy, sentence diversity, content relevance and workplace standardization, were statistically analyzed, and the results are shown in Table 2 below.

It can be seen from Table 2 that:

- Before the experiment, there was no significant difference in the scores of the four core dimensions of Workplace English writing between the experimental group and the control group ($P > 0.05$), indicating that the abilities of each dimension of the two groups of students were at the same level.
- After the experiment, the scores of the four dimensions of both groups improved, but the improvement range of the experimental group was significantly higher than that of the control group, and the inter-group differences all reached an extremely significant level ($P = 0.000 < 0.01$).

C. QUESTIONNAIRE RESULT ANALYSIS

After the experiment, a questionnaire survey was conducted on 58 students in the experimental group (effective recovery rate 96.67%), and the data were statistically analyzed using the Likert 5-point scale. The results are shown in Table 3 below.

It can be seen from Table 3 that the average overall evaluation score of the experimental group students on the deep learning-driven intelligent grading system is 4.14 points (full score 5 points), indicating that students have a high recognition of the system. Among them, 72.9% of students approve of the overall performance of the system (M≥4.0), showing that the application of the system meets students' learning needs and can effectively support students' workplace writing learning.

TABLE 1. Statistical Analysis of Comprehensive Writing Scores of the Experimental Group and the Control Group Before and After the Experiment (Unit: Points)

Group	Number of People	Test Type	Average Score (M)	Standard Deviation (SD)	t-value	P-value	Degree of Difference
Experimental Group	60	Pre-test	62.35	5.82	-18.76	0.000	Extremely Significant
Experimental Group	60	Post-test	78.62	4.95	-	-	-
Control Group	60	Pre-test	61.88	6.03	-8.32	0.001	Significant
Control Group	60	Post-test	68.95	5.77	-	-	-
Post-test Comparison Between Groups	-	-	-	-	9.25	0.000	Extremely Significant

TABLE 2. Statistical Analysis of Scores in Each Dimension of Writing Ability of the Experimental Group and the Control Group Before and After the Experiment (Unit: Points)

Group	Test Type	Grammatical Accuracy	Sentence Diversity	Content Relevance	Workplace Standardization
Experimental Group	Pre-test	14.22±2.18	13.85±2.12	15.12±2.25	13.16±2.23
Experimental Group	Post-test	20.35±1.95	19.78±1.88	20.56±2.01	18.93±1.92
Improvement Range of Experimental Group	-	6.13 (43.1%)	5.93 (42.8%)	5.44 (36.0%)	5.77 (43.9%)
Control Group	Pre-test	14.08±2.21	13.72±2.15	14.95±2.30	13.05±2.28
Control Group	Post-test	17.25±2.03	16.32±1.98	17.88±2.05	15.62±2.01
Improvement Range of Control Group	-	3.17 (22.5%)	2.60 (18.9%)	2.93 (19.6%)	2.57 (19.7%)
Post-test Difference t-value	-	7.65	7.23	6.89	7.92
Post-test Difference P-value	-	0.000	0.000	0.000	0.000
Post-test Difference Degree	-	Extremely Significant	Extremely Significant	Extremely Significant	Extremely Significant

TABLE 3. Statistical Results of Questionnaire Survey of Students in the Experimental Group (N=58)

Survey Dimension	Average Score (M)	Standard Deviation (SD)	Recognition Rate (Proportion of M≥4.0)
System Use Experience	4.23	0.68	77.6%
Writing Interest	4.15	0.72	72.4%
Autonomous Learning Ability	4.08	0.75	69.0%
Workplace Writing Confidence	4.12	0.70	70.7%
Overall Evaluation	4.14	0.65	72.9%

V. CONCLUSION

This study constructs an intelligent grading system suitable for higher vocational Workplace English writing with deep learning technology as the core, and draws the following core conclusions: First, the deep learning-driven intelligent grading system can significantly improve higher vocational students' comprehensive scores in Workplace English writing.

Empirical results show that the average comprehensive writing score of the experimental group students after the experiment increased to 78.62 points, with an increase rate of 26.1%, and there was an extremely significant difference with the post-test score of the control group ($P < 0.01$), which was much higher than the 11.4% increase rate of the control group.

This shows that compared with traditional teacher grading teaching, the intelligent grading system, with its advantages of instant feedback and accurate grading, can help students quickly correct writing errors, efficiently improve the comprehensive level of writing, and effectively solve the pain points of low grading efficiency and delayed feedback in traditional teaching, verifying research hypothesis H1.

Second, the intelligent grading system can comprehensively improve the four core dimensions of higher vocational students' Workplace English writing ability.

From the empirical data, the improvement range of the experimental group students in the four dimensions of grammatical accuracy, sentence diversity, content relevance and workplace standardization all exceeded 35%, among which the workplace standardization dimension had the highest improvement range (43.9%), and the post-test scores of the four dimensions were all extremely significantly different from those of the control group ($P < 0.01$).

This result shows that the system can not only solve students' basic errors such as grammar and spelling, but also targeted improve students' sentence application ability, content expression ability and workplace document specification mastery ability, fitting the "workplace-oriented" teaching goal of higher vocational English, verifying research hypothesis H2.

Third, the intelligent grading system has been widely recognized by students and teachers, and can effectively improve students' learning experience and literacy.

Questionnaire results show that the overall recognition rate of the experimental group students on the system reaches 72.9%, and the average scores of the four dimensions of system use experience, writing interest, autonomous learning ability and workplace writing confidence all exceed 4.0 points (full score 5 points); This shows that the intelligent grading system can not only improve students' writing ability, but also stimulate students' learning interest, cultivate autonomous learning ability, optimize teachers'

teaching mode, and verify research hypothesis H3.

AUTHOR CONTRIBUTIONS

Conceptualization –Hui Li and Yongmei Zhou; methodology – Hui Li and Yongmei Zhou; investigation –

Hui Li and Yongmei Zhou; writing-original draft preparation – Hui Li and Yongmei Zhou. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

HUMAN RESEARCH SUBJECTS

This study was approved by the Ethics Committee of Beijing Polytechnic College, and was performed in accordance with the principles of the Declaration of Helsinki. All eligible participants signed an informed consent form.

REFERENCES

[1] China Daily, "Vocational education reform made major national goal," 2023, [Online]. Available: https://english.www.gov.cn/policies/policywatch/202301/18/content_WS63c76385c6d0a757729e5c5a.html.

[2] J. Lyu, "Cultivating Cross-Cultural Competence in Students," SHS Web of Conferences, vol. 187, 2024, doi: 10.1051/shsconf/202418704006.

[3] R. Zhou, A. Samad, and T. Perinpasingam, "A systematic review of cross-cultural communicative competence in EFL teaching: insights from China," Humanities and Social Sciences Communications, vol. 11, no. 1, pp. 1750, 2024, doi: 10.1057/s41599-024-04071-5.

[4] M. Al-Badawi and A. Al-Tarawneh, "Cross-Cultural Communication Strategies for Business Professionals," in Frontiers of Human Centricity in the Artificial Intelligence-Driven Society 5.0. Reyad S, Hannoon A, Editors. 2024, Springer Nature Switzerland: Cham, pp. 1027-1032, doi: 10.1007/978-3-031-73545-5_92.

[5] L. Liu, "Research on Curriculum Construction of Workplace English in Higher Vocational Education," in Education and Educational Technology. 2012. Berlin, Heidelberg: Springer Berlin Heidelberg, doi: 10.1007/978-3-642-24775-0_127.

[6] Z. WeiWei, "Development and Practice of Work - style Vocational English Curriculum in Higher Vocational Education," In Proceedings of the 2017 International Conference on Innovations in Economic Management and Social Science (IEMSS 2017), pp. . Atlantis Press, 2017, doi: 10.2991/iemss-17.2017.183.

[7] H. Zhou, "Public English Teaching Strategies in Higher Vocational Colleges Based on Big Data Analysis of Students' English Test Scores," in Computer Science and Educational Informatization. 2024. Singapore: Springer Nature Singapore, doi: 10.1007/978-981-99-9492-2_28.

[8] A. B. Rashid and M. D. A. K. Kausik, "AI revolutionizing industries worldwide: A comprehensive overview of its diverse applications," Hybrid Advances, vol. 7, Art. no. 100277, 2024, doi: 10.1016/j.hybadv.2024.100277.

[9] D. Khurana, et al., "Natural language processing: state of the art, current trends and challenges," Multimed Tools Appl, vol. 82, no. 3, pp. 3713-3744, 2023, doi: 10.1007/s11042-022-13428-4.

[10] L. Y. Tan, S. Hu, D. J. Yeo, et al., "A Comprehensive Review on Automated Grading Systems in STEM Using AI Techniques," Mathematics, vol. 13, pp. 2828, 2025, doi: 10.3390/math13172828.

[11] S. Wang, G. Wang, X. Chen, et al., "A Review of Content Analysis on China Artificial Intelligence (AI) Education Policies," in Artificial Intelligence in Education and Teaching Assessment. 2022:1-8, doi: 10.1007/978-981-16-6502-8_1.

[12] X. Liu-Schuppener, "Artificial Intelligence and Digitalization in China's Education System," A Systematic Analysis of the Policy Framework and Underlying Strategies, 2023.

[13] Z. Y. Dong, et al., "Smart campus: definition, framework, technologies, and services," 2020; 2(1): 43-54, doi: 10.1049/iet-smc.2019.0072.

[14] A. Ayaan and K. W. Ng, "Automated grading using natural language processing and semantic analysis," MethodsX, vol. 14, Art. no. 103395, 2025, doi: 10.1016/j.mex.2025.103395.

[15] Rahanra N , Hossam A , Scott J ,et al.UTILIZING ARTIFICIAL INTELLIGENCE (AI) FOR AUTOMATED FEEDBACK ON THE ENGLISH ESSAY WRITING SKILLS OF INDONESIAN UNIVERSITY STUDENTS[J].JILTECH: Journal International of Lingua & Technology, 2025, 4(3), doi: 10.55849/jiltech.v4i3.1121.

[16] F. Branda, M. Ciccozzi, and F. Scarpa, "Artificial intelligence in scientific research: Challenges, opportunities and the imperative of a human-centric synergy," Journal of Informetrics, vol. 19, no. 4, Art. no. 101727, 2025, doi: 10.1016/j.joi.2025.101727.

[17] I. H. Sarker, "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions," SN Computer Science, vol. 2, no. 6, pp. 420, 2021, doi: 10.1007/s42979-021-00815-1.

[18] J. Wilson and A. Czik, "Automated essay evaluation software in English Language Arts classrooms: Effects on teacher feedback, student motivation, and writing quality," Computers & Education, vol. 100, pp. 94-109, 2016, doi: 10.1016/j.compedu.2016.05.004.

[19] X. Li and C. Huang, "Design of an intelligent grading system for college English translation based on big data technology," Systems and Soft Computing, vol. 7, Art. no. 200205, 2025, doi: 10.1016/j.sasc.2025.200205.

[20] M. Guo and Y. Li, "An Automatic Grading System for Chinese Primary School Students' Compositions," In Proceedings of the 2024 3rd International Conference on Artificial Intelligence and Education, pp. Association for Computing Machinery. p. 543-547, 2025, doi: 10.1145/3722237.3722332.

[21] X. Sui and L. Cheng, "Research on the Interactive Teaching Mode of College English Writing MOOC Assisted by Automatic Composition Marking System," Modern Educational Technology, vol. 29, no. 2, pp. 66-71, 2019, doi: 10.3969/j.issn.1009-8097.2019.02.010.

[22] J. Tang, "Research on the Application Mode of Automatic Writing Evaluation System in English Teaching," Foreign Language Teaching Theory and Practice, no. 1, Art. no. 9, 2014.