The Efficacy of Intelligent Evaluation and Peer Review in Improving the English Writing Ability of Vocational College Students

Weipeng Shen[1,*], Linhua Peng[2], Ruiqing Chen[1], Nan Jiang[1], and Xiao-Fan Lin[1,3]

[1] School of Information Technology in Education, South China Normal University, Guangzhou 510631, China
[2] Shenzhen Experimental Guangming School, Shenzhen 518107, China
[3] Guangdong Engineering Technologies Research Centre of Smart Learning, Guangzhou 510631, China
*Correspondence: swp784350292@gmail.com (W.S.)

Abstract—**The application of artificial intelligence has benefited vocational English writing courses but still lacks the capacity to dispose of contextual challenges. This study compared the effectiveness of two teaching methods in a vocational English writing course. A teaching method that combines intelligent evaluation and peer review was applied for the experimental group, while the control group adopted intelligent evaluation and conventional discussion. In this study, a quasi-experimental research method was used. Scores of students' essays as well as the quality of the content were examined. The results exhibited that the experimental group showed greater improvement than the control group. Thus, the designed teaching method provides a valuable method for English writing courses in vocational education.**

Keywords—**intelligent evaluation, peer review, English writing, vocational education**

I. INTRODUCTION

Under the background of international competition, the development of vocational education is gaining increasing attention. Vocational college students were facing challenges from the international market, which requires capable students to communicate, apply, and work with foreign languages. Hence, English writing courses in vocational colleges need renewal to cultivate students' requisite competency. The challenge requires efficient tools and methods to sufficiently cultivate students' writing skills. In vocational colleges, artificial intelligence has been widely applied in English writing courses for automated writing evaluation as intelligent evaluation. Intelligent evaluation usually includes overall ratings, grammar revisions, and written styles of students' English compositions, which helps students quickly and accurately analyze their compositions and provide revision advice [1].

Despite its advantages, intelligent evaluation in teaching platforms still has obvious limitations.

Specifically, intelligent evaluation was criticized because it can only provide limited feedback that only reminds students of grammar errors and basic word collection [2, 3]. But vocational college students' writing in complex workplaces relies on necessary situational skills, most of which still can not be assessed by intelligent evaluation systems. To generate students' situational skills, peer review has been proven to be an effective teaching method. In peer review activities, students need to assess the work or outcomes of their peers through scoring or verbalizing. Their reflection and critical thinking are accelerated in the process [4]. Students' critical reflection and communication further support students in re-conceptualizing, integrating, and creating past knowledge as well as experiences, thus developing situational skills [5]. It is possible that peer review can compensate for the shortcoming of intelligent evaluation and conventional discussion in vocational English writing courses. Therefore, this study combined intelligent evaluation and peer review in a teaching method designed for English writing courses in vocational colleges. The Results of the proposed method were validated and compared with the teaching method based on intelligent evaluation and conventional discussion to answer two research questions:

(1) Which teaching method (intelligent evaluation and peer review VS intelligent evaluation and conventional discussion) more effectively improves vocational college students' English writing performance?

(2) What are the differences between the result of the two teaching methods?

II. LITERATURE REVIEW

A. Vocational English Writing

Vocational education is an important part of the national education system and has the important responsibility of cultivating diverse talents, transmitting technical skills, and promoting employment and entrepreneurship [6]. The development of vocational education has received constant attention. Cultivating students' foundational capabilities in vocational

education is one of the important goals of vocational education [7]. English writing teaching is one of the foundation courses for vocational education students, which is related to their foreign communication and job-seeking ability. Therefore, it is imperative to adjust and optimize the vocational English courses.

B. Intelligent Evaluation in Teaching English Writing

Currently, vocational education benefits from the combination of Artificial Intelligence (AI) and education. Artificial intelligence is widely used in the writing and revision stages of English writing to provide learners with timely error-correcting feedback and assess the quality of their essays. Intelligent evaluation is a common method of applying artificial intelligence to teaching English writing in vocational education [8]. The intelligent evaluation produces timely feedback reports that can help teachers quickly and accurately analyze students' essays [9].

However, the current intelligent evaluation contains artificial intelligence techniques that can only analyze the grammar, semantics, and logic of students' English compositions [10]. Vocational students learn English writing mainly to solve problems in specific scenarios such as recruitment and foreign-related communication. Therefore, providing students with simulations close to real situations in writing instruction should be an important part of English writing in vocational education. However, no research has been conducted to further explore how to meet the needs of practical simulations for teaching English writing in vocational education based on the application of intelligent evaluation.

C. Peer Review in Teaching English Writing

Peer review is a teaching method widely used in teaching English writing that makes students engage in evaluating others' work by scoring or giving comments [11]. When students give and receive feedback in this process, peer review could effectively improve students' writing quality, leading them to develop their writing skills and think about the relationship between their writing and the situational workplace [12]. Therefore, peer review has been widely used in English writing courses. A previous study designed a quasi-experiment in which students in the experimental group were not only writers but also feedback providers with a guideline of questions for giving feedback, and the result showed that students with peer review presented better writing performance than those without peer review [13].

It is reported that when intelligent evaluation is accompanied by human assessment, they could improve students' writing skills by promoting meaningful discussions [14]. As peer review could also provide more systematic feedback and constructive advice while they focus on the structure, content, and language usage of peer's writing. Thus, it is possible that the implementation of both assistance in vocational English writing courses could help students improve their writing ability more efficiently.

III. MATERIALS AND METHODS

A. Participants

A total of 113 vocational students from two classes of vocational colleges in Guangzhou participated in this study. Two classes were matched for equal groups, with 51 students (28 male and 23 female) in the experimental group and 62 students (32 male and 30 female) in the control group. The average duration of English learning for all students was 10 years, and their basic abilities were similar.

B. Instruments

1) The standard for evaluating the English essays

The research revised the standard of English letter writing in the workplace, which was based on the second language writing standard of Jacobs *et al.* [15] and the English teaching request of higher vocational colleagues. The standard included four aspects: structure (25%), content (35%), form (20%), and language usage (20%), which were used in teachers' evaluations and the peer review of the students' writings.

2) The teaching platform

The Blue Ink Cloud classroom system was used to carry out students' peer review and intelligent feedback in this research, with the functions of checking the structure, word usage, grammar, and spelling in the students' writings with artificial intelligent evaluation. There are two main functions we used, intelligent evaluation and peer review, and the details can be seen in Figs. 1 and 2.
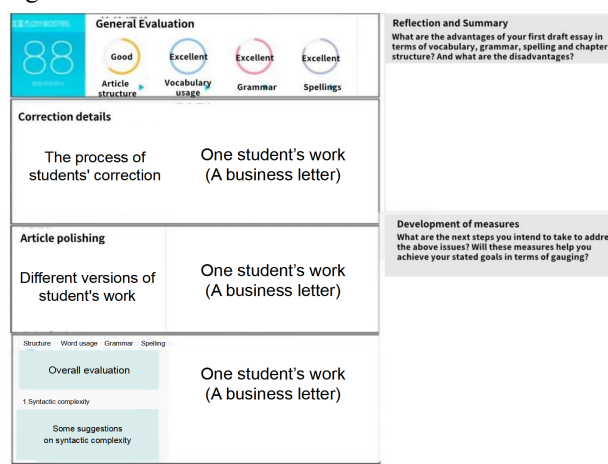

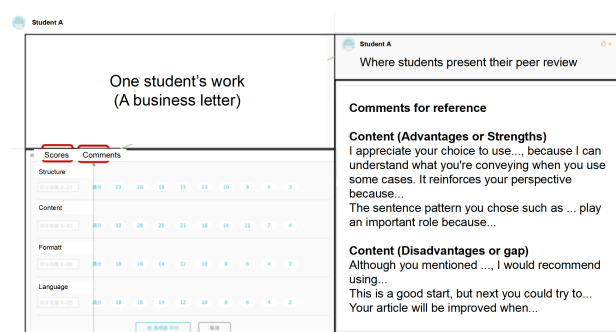Figure 1. The intelligent evaluation function of the Blue Ink Cloud classroom system.

Figure 2. The peer review function of the Blue Ink Cloud classroom system.

C. Research Design

A quasi-experimental research method was used in this study. The experiment lasted for six weeks, with two 40-minute sessions per week. Both groups were taught to use the Blue Ink Cloud class system and the process of peer review as pre-training in the first week. In the following 4 weeks, all students were taught to complete the tasks of contextual writing for real-life foreign language communication [16]. The only difference between the two groups was the use of peer review and conventional discussion. Students in the control group could choose to discuss online using the discussion function provided by the system, while students in the experimental group were guided with organized peer review. All of the students were tested on writing scores and writing quality for the pretest in week 1 and the post-test in week 6. The process of the experiment is shown in Fig. 3.
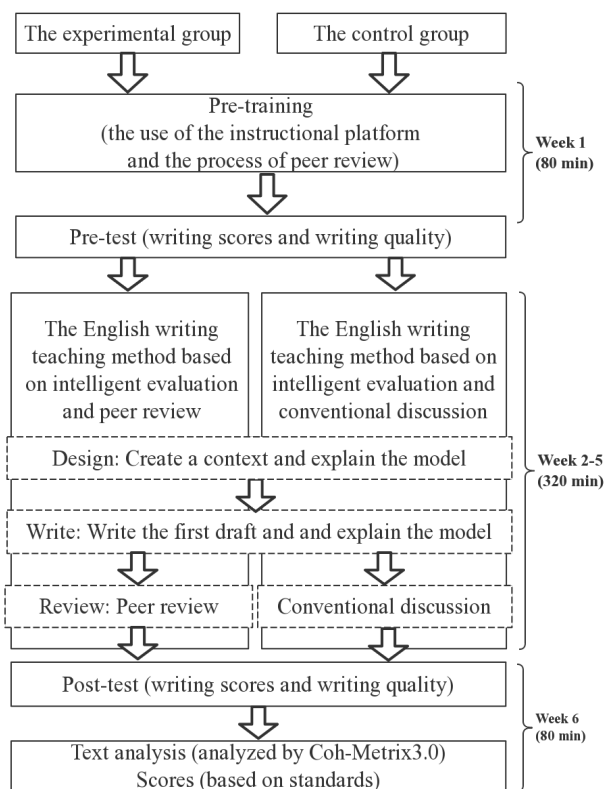

Figure 3. Research process diagram.

Fig. 4 shows the structure of the designed "intelligent evaluation and peer review teaching" method. From weeks 2 to 5, the teaching process was divided into three phases. In the design phase, students got contextual-writing tasks and learned the standard prepared for teachers and themselves to evaluate the letters. In the writing phase, students wrote drafts and uploaded them in the modified module of the Blue Ink Cloud classroom system. Students should repeatedly modify their drafts and upload them in the evaluation module to perfect their works with intelligent evaluation. Finally, students in the experimental group would use the review module of the platform to finish their peer review tasks in this phase. They would receive comments from different classmates to help them finish their letters, while students in the

control group could choose to discuss on the platform or not. After the final version of their work was uploaded, the result from the Blue Ink Cloud classroom system would contribute half of their final scores, and another half was decided by teachers. Students' final scores as well as their letters were collected for analysis.
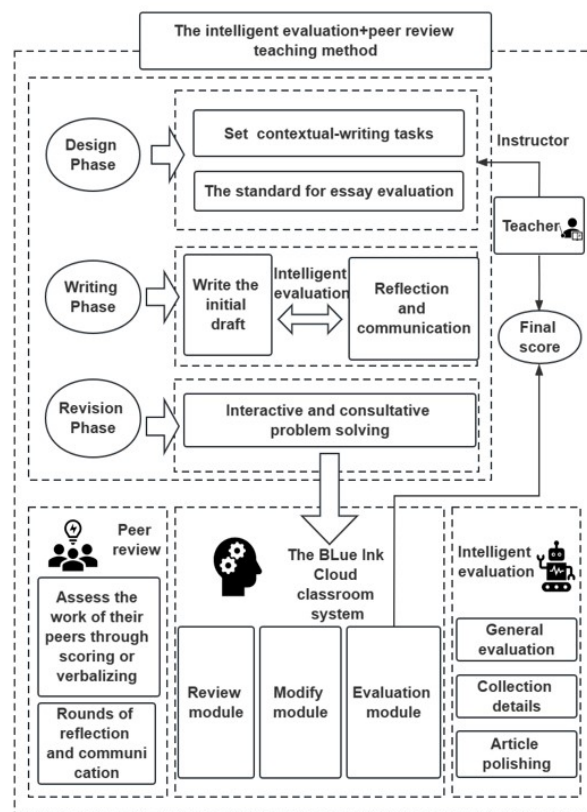

Figure 4. The structure of the designed intelligent evaluation and peer review teaching method.

D. Measures

Evaluation gauges of writing scores and the Coh-Metrix system (a text analysis system) were used to evaluate the influence of the method on students' writing quality. Every student's scores for the first draft and the final draft were collected from the platform automatically and the experimenter's evaluation. The average scores from these two ways were students' final scores.

Coh-Metrix is a computational system used to evaluate some textual features in written texts, which is widely applied to linguistic research studies. Additionally, the latest version includes 106 groups of variables of textual features like syntactic complexity. With Coh-Metrix, descriptive statistics, readability, and syntactic complexity were chosen to analyze the students' text.

IV. RESULT

A. Students' Achievements

The result of the ANCOVA is in Table I. Excluding the impact of the pretest and post-test, Table I shows that the adjusted mean of the experimental group of the final draft scores is 75.69 (SD = 7.97), while the control group is 69.16 (SD = 10.59).

TABLE I. THE ANCOVA ANALYSIS OF STUDENTS' WRITING SCORES

Group	N	Mean	SD	Adjusted mean	Std. error	F	η2
The experimental group	41	76.96	7.97	75.69	1.21	12.70***	0.216
The control group	50	67.59	10.59	69.16	1.35		

Note: *** $p < 0.001$.

B. Text Analysis

Coh-Metrix 3.0 and matched samples t-test were used to analyze text in the pretest and the post-test of the experimental group and the control group, and we learned the difference between writing quality in the two groups from the results. The results are shown in Tables II and III.

TABLE II. THE DIFFERENCE IN WRITTEN TEXTS BETWEEN THE PRETEST AND POST-TEST IN THE EXPERIMENTAL GROUP

Dimension	Variable	M	SD	t	Sig.
Descriptive	Paragraph length, number of sentences in a paragraph	−3.34	0.54	−6.17	0.00***
	Word length, number of syllables	−0.04	0.02	−2.07	0.04*
Text Easability Principle Component Scores	Text Easability PC Referential cohesion, z score	−0.73	0.29	−2.52	0.02*
	Text Easability PC Deep cohesion, z score	−0.20	0.22	−0.89	0.382
	Text Easability PC Verb cohesion, z score	−0.06	0.31	−0.21	0.838
Syntactic Complexity	Sentence syntax similarity, adjacent sentences	−0.01	0.02	−0.81	0.43
	Sentence syntax similarity, all combinations, across paragraphs, mean	−0.01	0.01	−1.30	0.21

Note: * $p < 0.01$, *** $p < 0.001$.

For descriptive statistics, both groups had significant differences that the average sum of words and syllables increased in the pretest and the post-test, and the mean of the post-test was higher than the mean of the pretest. Concerning feasibility, there was only a significant difference in referential cohesion Z, and the mean of it also increased. Additionally, in the experimental group, the means of verb cohesion Z and deep cohesion Z in the post-test were higher than those in the pretest. On the contrary, the means of the two variables above in the post-test were lower than those in the pretest in the control group. It was indicated that the improvement of text readability in the post-test was relatively obvious in the experimental group. As for syntactic complexity, both groups showed no significant differences in the pretest and post-test, but the means also increase in the experimental, while only the means of syntactic similarity adjacent sentences and all the sentences increased in the control group.

TABLE III. THE DIFFERENCE IN WRITTEN TEXTS BETWEEN THE PRETEST AND POST-TEST IN THE CONTROL GROUP

Dimension	Variable	M	SD	t	Sig.
Descriptive	Paragraph length, number of sentences in a paragraph	−2.76	1.72	−1.61	0.02*
	Word length, number of syllables	−0.53	0.02	−2.63	0.02*
Text Easability Principle Component Scores	Text Easability PC Referential cohesion, z score	−0.63	0.25	−2.56	0.02*
	Text Easability PC Deep cohesion, z score	−0.26	0.24	1.06	0.30
	Text Easability PC Verb cohesion, z score	0.07	0.29	0.23	0.82
Syntactic Complexity	Sentence syntax similarity, adjacent sentences	−0.02	0.02	−0.96	0.35
	Sentence syntax similarity, all combinations, across paragraphs, mean	−0.02	0.01	−1.53	0.14

Note: * $p < 0.01$.

V. CONCLUSION AND DISCUSSION

The study investigated how a teaching method combined with intelligent evaluation and peer review improves vocational college students' English writing performance. For question 1, the result indicates that the intelligent evaluation and peer review approach outperformed the intelligent evaluation and conventional discussion approach. For question 2, the result of students' scores confirms that peer review helps students get higher achievement in English writing tasks.

In the context of teaching English writing in vocational colleges, the study confirmed that peer review can compensate for the shortcomings of the conventional method. Intelligent evaluation was known to have difficulties in providing specific feedback in authentic situations [17]. Students can also simply get higher scores with more sophisticated vocabulary and longer essays [2]. By adding peer review, students can get more accurate advice on the organization and content of their writing [18], which may also make their writing more adaptable to real-life situations. Compared to the conventional method, the designed method enabled students to better understand the contextual-writing task in the process of reflecting on peers' comments and rewriting their compositions while receiving peer feedback, which is also in alignment with Pham *et al.* [19]. In the study, the result indicated that the addition of peer review can promote the content and structure of students' works [20]. Besides, though both the strategies used in the two groups improved students' writing scores and writing quality, the experimental group showed more improvement.

It is hoped that the designed method could compensate for the limitation of conventional intelligent English writing courses. We expected that the present study could contribute to designing effective teaching methods that

exploit the effectiveness of peer review for vocational English writing classes. It is also suggested that the optimization of the intelligent evaluation platform could be achieved by adding the peer review module. However, there are still some limitations in the research. Firstly, the method in this research is just used for teaching writing English letters, whether it is suitable for other content of courses should be studied in further research. Additionally, except for students' writing scores and text analysis, some scales for higher-order thinking can be considered.

CONFLICT OF INTEREST

The authors declare no conflict of interest. Data is available upon request from the first author. The data is not openly available elsewhere.

AUTHOR CONTRIBUTIONS

Weipeng Shen conceptualized the research idea; Linhua Peng conducted the experiment and analyzed the data; WeiPeng Shen, Ruiqing Chen and Nan Jiang wrote the original manuscript; Xiao-Fan Lin reviewed and edited the manuscript; all authors had approved the final version.

FUNDING

REFERENCES

[1] C. Strobl, *et al.*, "Digital support for academic writing: A review of technologies and pedagogies," *Computers & Education*, vol. 131, pp. 33–48, Apr. 2019. doi: 10.1016/j.compedu.2018.12.005

[2] I. I. Bejar, M. Flor, Y. Futagi, and C. Ramineni, "On the vulnerability of automated scoring to construct-irrelevant response strategies (CIRS): An illustration," *Assessing Writing*, vol. 22, pp. 48–59, Oct. 2014. doi: 10.1016/j.asw.2014.06.001

[3] X. Lu, "An empirical study on the artificial intelligence writing evaluation system in China CET," *Big Data*, vol. 7, May 2019. doi: 10.1089/big.2018.0151

[4] S. Cao, S. Zhou, Y. Luo, *et al.*, "A review of the ESL/EFL learners' gains from online peer feedback on English writing," *Frontiers in Psychology*, vol. 13, Oct. 2022. doi: 10.3389/fpsyg.2022.1035803

[5] G. Ion, A. S. Martí, and I. A. Morell, "Giving or receiving feedback: Which is more beneficial to students' learning?" *Assessment & Evaluation in Higher Education*, vol. 44, no. 1, pp. 124–138, Jan. 2019. doi: 10.1080/02602938.2018.1484881

[6] E. Xue and J. Li, "Exploring the type-based vocational education system: Insights from China," *Educational Philosophy and Theory*, vol. 54, no. 10, pp. 1670–1680, Jun. 2020. doi: 10.1080/00131857.2021.1934668

[7] T. C. Hsu, W. L. Chen, and G. J. Hwang, "Impacts of interactions between peer assessment and learning styles on students' mobile learning achievements and motivations in vocational design certification courses," *Interactive Learning Environments*, vol. 5, pp. 1–13, Oct. 2020. doi: 10.1080/10494820.2020.1833351

[8] X. Cai, "Practice of hybrid teaching mode of English writing based on artificial intelligence," *Journal of Physics: Conference Series*, vol. 1648, no. 4, 042062, Oct. 2020. doi: 10.1088/1742-6596/1648/4/042062

[9] J. F. Chen and C. A. Warden, "Application of artificial intelligence to the small open online English abstract writing course," in *International Conference on Innovative Technologies and Learning*, Dec. 2019, pp. 802–808. doi:10.1007/978-3-030-35343-8_83

[10] Z. Wang, "Computer-assisted EFL writing and evaluations based on artificial intelligence: A case from a college reading and writing course," *Library Hi Tech*, vol. 40, pp. 80–97, Oct. 2020. doi: 10.1108/LHT-05-2020-0113

[11] Y. Wang, X. Luo, C.-C. Liu, *et al.*, "An Integrated automatic writing evaluation and SVVR approach to improve students' EFL writing performance," *Sustainability*, vol. 14, 11586, Sep. 2022. doi: 10.3390/su141811586

[12] K. Topping, "Peer assessment between students in colleges and universities," *Review of Educational Research*, vol. 68, no. 3, pp. 249–276, 1998. doi: 10.3102/00346543068003249

[13] J. Novakovich, "Fostering critical thinking and reflection through blog-mediated peer feedback," *Journal of Computer Assisted Learning*, vol. 32, no. 1, pp. 16–30, Oct. 2016. doi: 10.1111/jcal.12114

[14] A. Liaqat, C. Munteanu, and C. D. Epp, "Collaborating with mature English language learners to combine peer and automated feedback: A user-centered approach to designing writing support," *International Journal of Artificial Intelligence in Education*, vol. 31, pp. 638–679, July 2021. doi: 10.1007/s40593-020-00204-4

[15] S. Sujito, E. Yunita, *et al.*, "Applying direct peer feedback to foster vocational school students' English writing performance," in *Proc. Annual Conference on Social Sciences and Humanities*, 2018.

[16] H. Jacobs, S. Zinkgraf, D. Wormuth, *et al.*, *Testing ESL Composition: A Practical Approach*, Newbury House Publishers, Inc., Jan. 1981.

[17] K. Johnstone, H. Ashbaugh, and T. Warfield, "Effects of repeated practice and contextual-writing experiences on college students' writing skills," *Journal of Educational Psychology*, vol. 94, pp. 305–315, Jun. 2002. doi: 10.1037/0022-0663.94.2.305

[18] G. M. Awada and N. M. Diab, "Effect of online peer review versus face-to-face peer review on argumentative writing achievement of EFL learners," *Computer Assisted Language Learning*, vol. 36, pp. 238–256, Jun. 2021. doi: 10.1080/09588221.2021.1912104

[19] T. N. Pham, M. Lin, V. Q. Trinh, and L. T. P. Bui, "Electronic peer feedback, EFL academic writing and reflective thinking: Evidence from a Confucian context," *SAGE Open*, vol. 10, no. 1, 215824402091455, Jan. 2020. doi: 10.1177/2158244020914554

[20] H.-C. Tai, W.-C. Lin, and S. Yang, "Exploring the effects of peer review and teachers' corrective feedback on EFL students' online writing performance," *Journal of Educational Computing Research*, vol. 53, Aug. 2015. doi: 10.1177/0735633115597490