



Effect of Peer Assessment and Artificial Intelligence Review on English Writing and Evaluation Ability

Hong Lu

School of Foreign Languages, Zhengzhou University of Science and Technology, China, kiki1314ll@163.com

Peer assessment has been widely adopted in teaching English writing; artificial intelligence (AI) assisted teaching has also been practiced and applied in English writing teaching. Based on an in-class writing practice of 250 Chinese undergraduate students majoring in English and Translation, this study adopts an empirical comparative method and a questionnaire survey to carry out post-writing peer assessment and AI review. By comparing the similarities and differences through the practice of the two different modes, this study tries to explore their value in effectively improving undergraduates' English writing and evaluation abilities. The questionnaire survey shows that both modes can give constructive evaluations on the content; organization; sentence expression; use of words; grammar, spelling, punctuation and other technical aspects of an essay, with the AI review providing more objective and specific feedback. In addition, in the artificial intelligence era, the use of peer assessment and AI review can optimize English writing teaching and enhance the ability of undergraduates' English writing and evaluation.

Keywords: peer assessment, artificial intelligence (AI), AI review, empirical study, English writing ability, evaluation ability, EFL

INTRODUCTION

Nowadays, English writing ability is important for the development of professionals, and writing is the weakest ability in students' English skills (Huang Yuanshen, 2006), which is difficult to make significant progress. Developing writing ability has essential advantages of foreign language learners (Benlaghrissi & Ouahidi, 2024). Moreover, the learning of English writing transcends mere skill enhancement; it also involves the ability to critically evaluate one's own work and that of peers. Evaluation ability are not only crucial for effective teaching but also play a pivotal role in fostering students' writing ability by promoting reflective learning. And with the in-depth development of foreign language education in China, the main body of educational evaluation has tended to be diversified, not only teachers but also students need to have the ability to evaluate. Therefore, students' English writing and evaluation ability should be improved. In particular, the improvement of students' evaluation ability can promote the

Citation: Lu, H.. (2025). Effect of peer assessment and artificial intelligence review on English writing and evaluation ability. *International Journal of Instruction*, 18(4), 517-530.

occurrence of deep learning (Torshizi & Bahraman, 2019). So such tools as AI should be used to assist teaching and learning in the context of the general environment in the age of artificial intelligence.

The traditional model of teaching is teacher-centered and students have high trust in teachers to evaluate their writing. Teachers are usually regarded as the key to writing evaluation (Widiastuti, 2021), and writing evaluation is also an unavoidable challenge for teachers (Ferris and Hedgcock, 2014), which is burdensome. Azmar and Razali view (2024) that teaching writing is a tedious task. So as an effective supplement to teacher evaluation, peer assessment is popular in English writing teaching. Peer assessment emphasizes students' central subjectivity, which can motivate students to participate in writing evaluation. This mode can reflect on students' own writing practice and foster active revision of their compositions. In the rapid integration of AI and the field of education, AI is crucial for educators (Gupta et al., 2024; Ng et al., 2021; Wang & Wang, 2025) and as an effective tool to assist English writing teaching, can also help improve English writing ability. In particular, AI review is flexible and convenient, can provide instant feedback, and is not limited by time and place, so it can act as an "assistant" in the process of English writing and evaluation, and empower the improvement of English writing and evaluation ability.

In the continuous exploration of English writing teaching and learning, attempts are made to organically combine and complement different evaluation modes, such as teacher evaluation, peer assessment and AI review. Those attempts aim to find the most suitable feedback path for undergraduate students' post-evaluation and improvement of English writing. There are few studies focus on peer and AI evaluations to find the development of students in their writing processes (Lisa et al., 2025). Meanwhile, given the challenges associated with traditional writing teaching and the potential benefits of integrating AI into the evaluation process, this study aims to investigate the impact of peer assessment and AI review on undergraduate students' English writing and evaluation ability. Specifically, this study seeks to find the differences between these two evaluation modes and how they can contribute to the development of students' writing and evaluation ability, and tries to provide concrete suggestions in using them during English writing teaching and learning.

All in all, this paper will focus on the influence of peer assessment and AI review on undergraduates' English writing and evaluation ability. The findings of this study are expected to have significant implications for English language learners and educators, as they seek to optimize their learning and teaching methods

Current Situation

The object of this study is the undergraduate students majoring in English and Translation, who are offered a course of *English Writing I* in the second semester of their freshman year and *English Writing II* in the first semester of their sophomore year. *English Writing* is the foundation course for their majors. The textbook used in the two semesters is *A Handbook of Writing* published by Foreign Language Teaching and Research Press and edited by DingWangdao, et al. The teaching content of *English Writing I* (Stage I) focuses on students' understanding and using of English writing

format, dictions, sentences and paragraphs, while *English Writing II* (Stage II) involves the writings of different topics such as descriptive, narrative, expository and argumentative essays, and different types of writing such as research papers and practical essays. This study will analyze the current situation of English writing evaluation based on Stage I, and conduct peer assessment and AI review based on Stage II.

In Stage I, the teacher had assigned three writing assignments. The first task focuses on checking students' mastery of writing format, the second one is a paper form of paragraph writing on a given topic, and the third one is an electronic version of paragraph writing. All these three practices were evaluated by the teacher who spent more time reviewing and reflecting. Jacobs et al. (1998) found that students valued teacher feedback in particular because of the influence of a teacher-centered culture. Teachers' comments for a writing have a certain value for students, especially content analysis and encouragement for students, but the perspective of teacher is affected by individual preferences and is subjective. Secondly, the time between the completions of writing assignments to the feedback should not be too long, otherwise it will not be conducive to the reinforcement of the students' thinking. That requires teachers to complete the evaluation at a high level of intensity, and the workload is heavier.

Stage II has a greater integrity in terms of the content focusing on composition writing. This also raises the requirements for homework evaluation, which is showed in the length and difficulty of the text. Therefore, in this stage, the teacher tries to use peer assessment and AI review to make up for the singularity of teacher evaluation. Before proceeding with the writing, the teacher introduced these two modes and their functions, i.e. peer assessment is an educational activity of mutual assessment, in which students exchange, read and give revisions to each other's writing (Mangelsdorf, 1992). In this writing practice, each student will review a copy of the other student's writing and give his or her own grades and comments. Meanwhile, AI can play an important role in teaching English writing, and it can be used in the generation of model essays and writing evaluation. In contrast to peer assessment, AI review will also be used into this writing practice to help students find a suitable way of evaluating for themselves and to enhance their subjective awareness. Before the practice, students generally have higher expectations of peer assessment, and are mostly in a state of not knowing or not having used AI review in English writing.

Combining the above analysis of the current situation, this study addresses three research questions. First, what are the differences between peer assessment and AI review? Second, how do these two modes perform in terms of grading and commenting on English compositions? Third, can these two modes, particularly the AI review, promote the ability of students' English writing and evaluation?

METHOD

Participants and procedure

This empirical study is based on Stage II, in which there are 250 students majoring in English and translation, distributed in 10 administrative classes and 5 teaching classes,

in the form of co-teaching. The specific process is as follows: students complete an essay by the given topic *Foreign Languages Should Be Taught in Nurseries or Kindergartens* within 30 minutes, and the word count is between 180 and 250 words. They can refer to their textbook but there is no use to mobile phones. Subsequently, 250 texts were obtained from this writing practice. Then the teacher instructed the students to complete peer assessment. The five reference items for students to use when revising their peer's composition are from *A Handbook of Writing*.

The five reference items are: content; organization; sentence; diction; grammar, spelling, punctuation and other technical aspects. Each reference item is worth 2 points, and the total grade of the essay is 10 points. Students are required to grade other students' essays against these 5 reference items, and a specific grade and comments are given at the end. In addition to the grade and comments, the teacher also set up an incentive, that is, students who think that the grade and comments they gave in this peer assessment are constructive can write down their names, which can get additional grade from the teacher.

Data collection and Analysis

The following are the statistics of peer assessment by the unit of teaching classes:

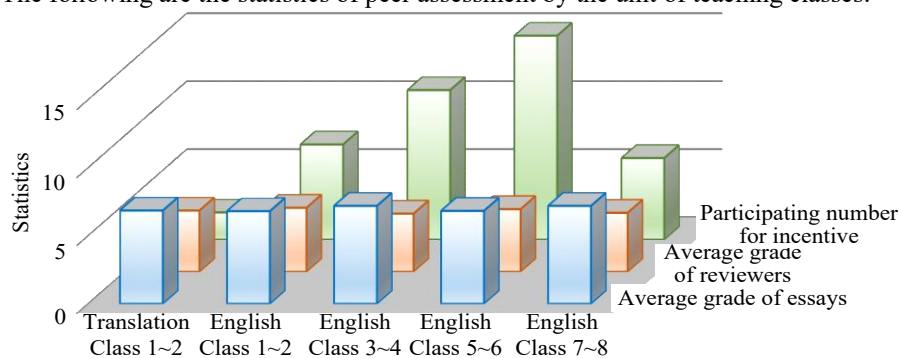


Figure 1
Statistics of peer assessment among average grade of essays and reviewers, and participating number for incentive

It can be seen that the highest average grade was 7.20 and the lowest was 6.81; the highest average grade given by the teacher to the reviewers was 4.70 and the lowest was 4.26; and the highest number in the incentive was 15, and the lowest was 2. While the differences in the first two are less significant, the last set of data is extremely variable. Moreover, the number of all students participating in the incentive is only 16.4% of the total, indicating that students have low confidence and self-identity in conducting peer assessment.

The second round is AI review based on ERNIE Bot 3.5. Firstly, the teacher presents a sample essay on the same topic in the form of a dialogue with ERNIE Bot 3.5, and also shows the AI grade and comments of the given sample essay. The reference items for the sample are consistent with the peer assessment. Secondly, in each teaching class, the

teacher selects one essay to present based on peer assessment grades and comments quality, and inputs it into ERNIE Bot 3.5. Subsequently, a comparative analysis was conducted on the results of AI review and peer assessment, guiding students to reflect on the similarities and differences between them, and the advantages and disadvantages of AI tools in English writing and evaluation. Here are the grades and comments of one sample in two evaluation modes:

Table 1
Grades and comments sample

Mode	Grade	Comments
Peer assessment	7.2	Content: rich; structure: reasonably well organized and appropriate; sentences: focused and more varied in type; diction: more rigorously worded; a few grammatical errors but not detracting from the full meaning of the essay.
AI review	6.0	<p>The essay basically expresses the author's opinion that kindergartens or nurseries should offer foreign language courses in terms of content, but there are some problems of unclear expression and lack of logic. For example, the sentence "Leaning language for children is more earlier, more better." has problems in grammar and expression, and should be changed to "Learning languages earlier is better for children." In addition, the article jumps around a bit in the course of the discussion, lacking the necessary transitions and connections.</p> <p>In terms of organization, although the essay has a certain structure, it is not compact and coherent enough on the whole. The beginning is rather abrupt and does not introduce the theme well; the middle part lists some reasons but lacks in-depth analyses and examples; and the ending is rather hasty and does not summarize the whole text well.</p> <p>In terms of sentence expression, the essay contains some grammatical errors and unclear expressions. For example, the sentence "There are a sea of time that they can touch foreign language." should be changed to "They have plenty of time to learn a foreign language." In addition, the sentence structure of the essay is rather homogeneous and lacks variety, which makes the overall expression rather bland.</p> <p>In terms of diction, the vocabulary in the essay is generally in line with the subject matter, but some of the words used are not accurate or are too simple. For example, the expression "more better" is incorrect and should be replaced by "better"; "the rapid development of society" should be replaced by "the rapid development of society".</p> <p>Apart from the grammatical errors mentioned above, there are some spelling mistakes and improper use of punctuation. For example, the sentence "And many parents have seen that many parents' families have a higher education make children receive education at their young." In this sentence, "has seen" should be changed to "have seen", "family" should be changed to "families", and "to" should be added before "make".</p>

After the teacher's demonstration, students are allowed to perform real-time field operations, input their own essay into the AI review system, get grade and comments, and compare the differences with peer assessment. Finally, students are encouraged to use writing assistant tools effectively, pay attention to the writing evaluation, and grow in the process of reflection.

Questionnaire Analysis

After the above demonstration of AI review and the comparison and reflection of the two modes of evaluation, students were asked to fill out a questionnaire, which consists ten questions, including eight single choices, one multiple choice, and one open-ended question. The questions cover the satisfaction with AI review and peer assessment, and the usefulness of both on the improvement of writing, etc. In the reliability of the

questionnaire, the Cronbach. α coefficient is 0.722 for the three questions related to peer assessment, 0.760 for AI review, and 0.842 for comparison of peer assessment and AI review. The reliability coefficients are all above 0.7, so the questionnaire results are valid.

As far as the satisfaction with AI review, most students are satisfied with 66.41% choosing “satisfied” and 11.72% “very satisfied”, which indicates that AI review has been widely recognized by students in writing evaluation. Secondly, regarding the degree of the help of AI review on the improvement of writing ability, 67.19% of the students choose “helpful” and 19.53% select “very helpful”. That shows the positive effect of AI review in improving writing ability. At the level of writing confidence, 81.25% feel that their writing confidence increased after using AI review, which further supports the positive impact of AI review.

In addition, the satisfaction level of peer assessment was lower than that of AI review, with only 51.56% indicating “satisfied”, which may reflect the inconsistency in the quality of feedback from peer assessment. Only 7.81% views “very helpful” and 52.34% “helpful”, indicating that the feedback from peer assessment can help improve writing skills to a certain extent, but the effect is not as significant as that of AI review. However, 68.75% of the students believe that peer assessment could provide more constructive feedback, showing the value of interpersonal interaction in evaluation. Moreover, in the comparison of the effectiveness of AI review and peer assessment, 60.94% believe that both are effective, while 34.38% believe that AI review is more effective and only 3.91% choose peer assessment.

As for the preference of the two modes, 55.47% of the students are willing to accept both AI review and peer assessment, and 37.5% prefer AI review, indicating that AI review is more popular among the participants. Moreover, students generally believe that AI review is more objective (82.81%) and its feedback is faster (71.09%), while peer assessment is more humane (72.66%). This reflects that participants have a clear understanding of the different characteristics of the two evaluation methods. At last, in response to the open-ended question “Briefly describe your overall feelings and suggestions after using AI review and participating in peer assessment”, the top three high-frequency words about peer assessment are humanize, warm, and so-so; for AI review, they are objective, comprehensive and fast.

In summary, AI review performs better than peer assessment in improving writing ability and confidence, and students’ acceptance and satisfaction with AI review was higher. Although peer assessment has its advantages in providing constructive feedback, the overall effectiveness and satisfaction still need to be improved.

FINDINGS

After the completion of the above practice, the author input the collected 250 pieces of corpus into ERNIE Bot 3.5 for AI review to get all the grades and comments.

Grade Comparison between Peer Assessment and AI Review

The following is the scatter diagram of the grade comparison between peer assessment and AI review: Figure (a) is a summary diagram of the ratio comparison between the

two modes, and Figure (b) to Figure (f) shows the grade comparison of the five teaching classes.

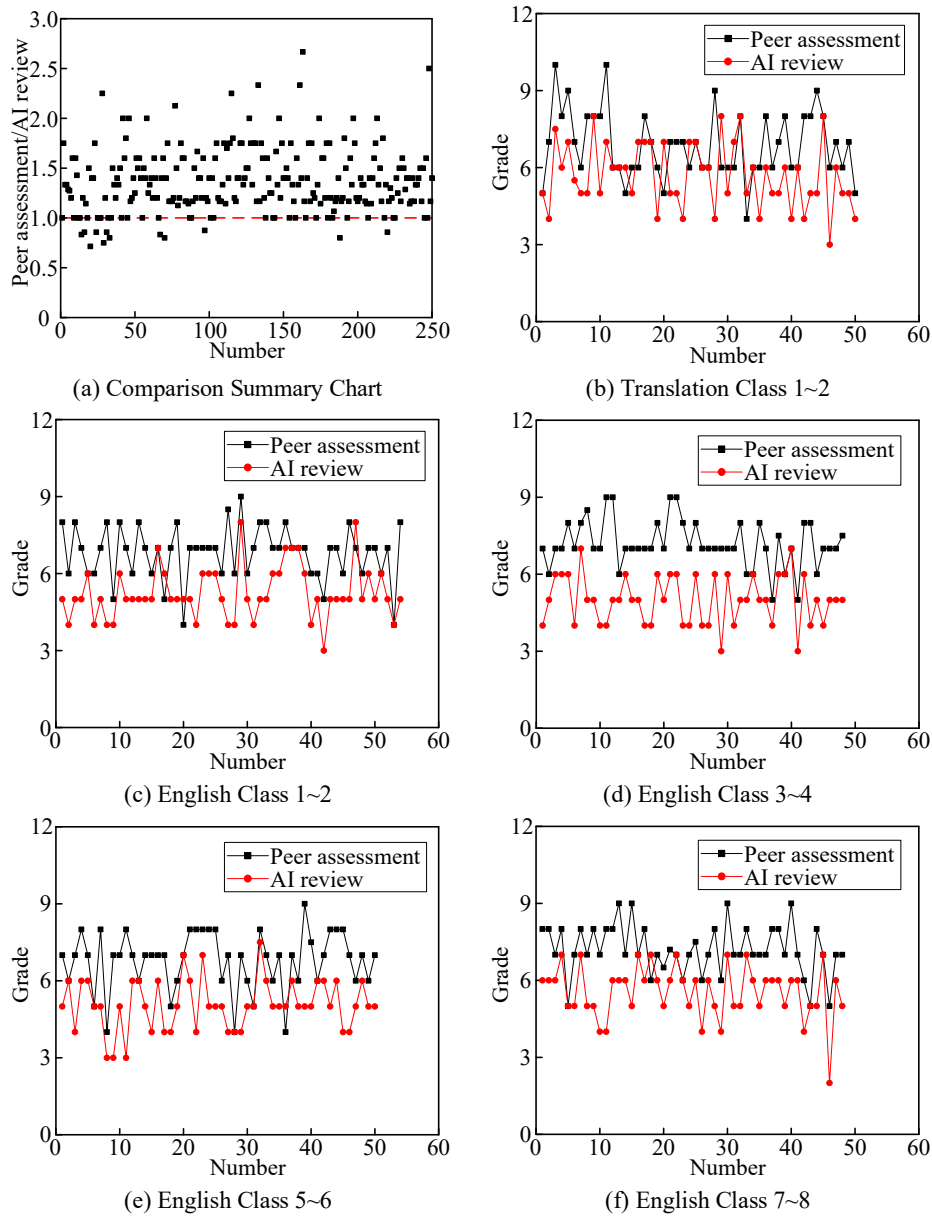


Figure 2
Scatter diagram of the grade comparison between peer assessment and AI review

From Figure 2-(a), it can be seen that peer assessment gives significantly higher grades than AI review. When broken down into teaching classes, overall most students' AI review grades are significantly lower than peer assessment, a few are equal, and there are also differences between different teaching classes. For example, English Class 3~4 have the greatest difference in the grades derived from the two modes of evaluation.

Comments Comparison between Peer Assessment and AI Review

In addition to the grades, the difference in comments between peer assessment and AI review is relatively large. The total number of words in the comments of the 250 corpus in peer assessment are 6,828 words, with an average of 27 words. However, the total number of words in the comments of AI review is 187,144 words, with an average of 749 words. Specifically, the comments of peer assessment are relatively simple, except for pointing out some technical errors such as grammar and spelling, and it is mostly encouraging, which can provide a certain emotional value to the peers. The comments of AI review, on the other hand, is strictly based on the five reference items of grading, pointing out the errors of the essays one by one as well as the optimization suggestions, which is more detailed and targeted, and has a high value of reference for the subsequent modification of the essays.

Considerations

Comparing the differences in grades and comments between the two evaluation modes, this study first considers the factors that affect the grades and quality of peer assessment. First, the peers' understanding of the evaluation criteria is inconsistent (Ma Xiaomei, 2023). There are limited resources and guidance available for reference during the evaluation, the evaluators' own evaluation ability varies, and the fairness and objectivity of the evaluation is affected by human relationships (Price et al., 2011). For example, one of the corpus essays obtains 10 points in the peer assessment and the comments are only "very good". Later, it was found that the writer of this essay is a member of the study committee who is popular among students, and the peer assessment grades and comments are not objective enough in the process. In the AI review, this essay is graded 7 out of 10, and the comments points out that the essay needed to be improved in its content, organization, sentence structure, use of diction and technical issues, and gives specific suggestions for revision.

What is more, Carson & Nelson (1996) found that the primary goal of Chinese students when engaging in peer assessment is to remain harmony. They are reluctant to criticize their peers' essays or disagree with peers' opinions, and are unwilling to take on an authoritative role. From the above study, most of the comments in the peer assessment are positively orientated, praising the excellent performance of their classmates' essays with less detailed descriptions of the substantive problems. And the comments mostly ended with encouraging words, such as "great", "Come on", etc. Next, according to the participation of the incentive, the ratings of the participants are positively related to the number of words in the comments, and most of them contains positive words. On the other hand, the peer comments of those who did not participate in the incentive are more likely to point out problems in content and sentence structure, thus countering some students to avoid participating the incentives by pointing out specific writing

problems of their classmates in peer assessment. This tendency can affect the authenticity and effectiveness of peer assessment. Comparing the comments of AI review, there is a substantial increase in the word count, the comments are closely related to the reference items, and the logic is clear and detailed.

Although peer assessment and automatic grading systems only point out language errors (Hanjani, 2016; Khonbi & Sadeghi, 2012; Tsui & Ng, 2000), it is difficult to make an in-depth evaluation on the discourse structure and logic of the essay and provide practical suggestions. However, for undergraduate students who are in the stage of consolidating the foundation of English writing, the improvement of language usage and content is crucial. There are researches shown that AI can provide high enough quality feedbacks in writing evaluation of college English teaching (Steiss et al., 2024). While the target of this study, English and Translation majors at undergraduate level, still show a lack of solidity at the language level, and other non-major undergraduates need to be evaluated after writing, which helps them to improve and enhance their English writing.

DISCUSSION

Evaluation is an important part of education (Xu & Brown, 2016) and an indispensable process for the improvement of undergraduates' English writing ability. Intervening peer assessment and AI review into the teaching of English writing courses, especially the evaluation after writing practice, teachers can ease the burden of homework correction and students can get more objective suggestions. Taking the teacher evaluation as an example, when assigning a writing assignment, the teacher needs to spend twice as much time to evaluate. However, with the integration of peer assessment and AI review, the feedback on grammar, spelling and punctuation can be got instantly, which improves teaching efficiency to a large extent. For undergraduate students, both modes are able to give useful feedback on the content; organization; sentence expression; use of words; grammar, spelling, punctuation and other technical aspects of an essay. In the context of this empirical study, peer assessment and AI review have their own strengths and are recommended to be used in combination.

IMPLICATIONS

For teachers, they should adapt to the AI era. Combining peer assessment and AI review to form the workflow of teaching English writing, especially exploring and optimizing peer assessment based on the automatic evaluation system, which is also a need for the current reform of foreign language writing teaching in China (Gao Ying, 2019). That is, in each writing practice, the teacher can start from determining the writing assignment (limiting the time and place, prohibiting the use of mobile phones, and clarifying the evaluation standards and grading rules) to complete the writing practice, conduct peer assessment (collecting peer assessment grades and comments), then conduct the AI review with a focus on comparing the results of the two modes of evaluation, and finally combine with questionnaires to learn about students' evaluation experiences. This workflow will enable more students to understand the two modes of English writing evaluation, and to use them to improve their own English writing and evaluation ability.

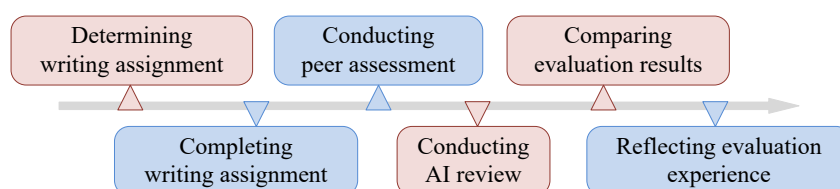


Figure 3
English writing teaching workflow

Next, teachers should provide more guidance to students before conducting peer assessment. It is known that students can participate in peer assessment extensively, but there are differences in students' English writing levels, which prevent them from evaluating and giving feedback on the problems of others' essays. In order to increase the effectiveness of peer assessment, each writing practice needs to be carefully planned, with clear evaluation standards, modelling and guidance. In the future use of peer assessment, teachers should strengthen the training before peer assessment, which can improve the efficiency and effectiveness of peer assessment. Moreover, teachers can enrich the forms of peer assessment, such as exchanging essays between different teaching classes, discussing peer assessment in small groups, and refining the standards of peer assessment to promote fairness.

Furthermore, teachers should pay more attention to students' perception of AI review. The guidance of teachers plays a crucial role in utilizing AI tools for learning foreign languages of students (Stevenson & Phakiti, 2014). Meanwhile, teachers need to seek strategies to promote students' perceptions in order to facilitate their output of adopting and practicing evaluation suggestions in writing revision. Then teachers can train students to be able to receive AI review comments objectively to avoid a blow to their writing confidence, so that they can reap more benefits from AI review. At the same time, students are guided to focus on enhancing assignment definiteness and target-object identity certainty when using AI assisted writing and evaluation through demonstration or modeling. In particular, let students know that the more targeted the directive, the stronger the role of AI assistance.

For students, they should update their concept and mode of English writing and evaluation. Students are well aware of the advantage of excellent English writing ability, but they lack a certain amount of writing practices in the first place, and not getting effective evaluation. That will not only dampen their confidence in writing, but also be detrimental to the improvement of writing ability. Most students especially want one-on-one feedback from their teachers after writing, but a teacher's evaluating perspective carries a certain degree of subjectivity. Multi-dimensional review of the same essay, such as adding peer assessment and AI review, can find more problems and increase the efficiency of revision.

In addition, it is important for students to incorporate revision suggestions. Writing ability is not only honed in writing process, but also refined in guided revision (Sun Shuguang, 2019). Not only teacher evaluation, but also peer assessment and AI review can provide valuable suggestions for students to revise their writing. High-quality

feedback or comments can facilitate learners' critical reflection (Liu & Carless, 2006). Therefore, the feedback and comments given by different evaluation methods should be revised and polished against the original essay carefully. Taking AI review as an example, students can continue to use AI review after revising according to the comments, and make several revisions until a satisfactory final draft is formed.

Finally, students need to realize the value of AI review but not misuse AI (Bowen & Watson, 2024). AI has a wide range of application scenarios and plays an important role in education (Wu Li, 2024). Undergraduate students are less restricted and more operational in the use of AI tools than students in other lower educational levels. In terms of English writing ability, undergraduates have a certain foundation and have received appropriate instruction in writing standards, and the AI review is suitable for use in English writing practice at the university level. Based on this study, non-English majors can also learn more about AI review on English writing through English courses, English campus activities or lectures, so as to encourage more college students to use AI to enhance their confidence in English writing and improve their ability to evaluate English writing.

CONCLUSION

Combined with this empirical comparative study and questionnaire survey on peer assessment and AI review of English writing, it shows that students' recognition of peer assessment and AI review is relatively high. In particular, AI as a new type of technological tool broadens the research perspective of English writing evaluation, and has strong advantages in improving students' English writing and evaluating abilities. It can provide real-time objective suggestions for students' English writing and evaluation. However, AI still has machine-like and limitation on the emotional value of conducting directive English composition evaluation while peer assessment shows more emotional communication and humanistic care.

Assessment is crucial for improving language proficiency (Díaz et al., 2023). So as for future teaching and learning of English writing, the practice of English writing evaluation should be actively promoted, and various evaluation modes such as teacher evaluation, peer assessment and AI review should be comprehensively used. Teachers should guide students to use AI efficiently in English writing content; organization; sentence expression; use of words; grammar, spelling, punctuation and other technical aspects. Writing logic and ideas should be listened to peers and teachers. Importantly, both should not rely on any one mode solely, especially AI, so as to strengthen the revision and reflection of English writing based on the feedback from multiple evaluating perspectives. All in all, peer assessment and AI review will do more benefits in cultivating students' critical writing and evaluating thinking in a deeper way and further improving their English writing and evaluation ability.

However, due to the limited number of students involved in this study and the fact that only one round of writing evaluation was conducted, there are certain limitations. Next, the research focus could be shifted to the quality of students' composition revisions after combining the two evaluation modes (peer assessment and AI review). Meanwhile, the research subject could also be centered on teachers to understand the questions and

findings they encounter when using and teaching these two modes. The above attempts are aimed to better promote evaluation-driven learning and teaching.

REFERENCES

- Azmar, K., & Razali, A. B. (2024). Teaching and learning of English academic writing in a Public University in Malaysia: A case study. *International Journal of Instruction*, 17(4), 645-668. <https://doi.org/10.29333/iji.2024.17436a>
- Benlaghissi, H., & Ouahidi, L. M. (2024). Effect of WhatsApp-based tasks on developing EFL students' writing skills. *International Journal of Instruction*, 17(2), 707-720. <https://doi.org/10.29333/iji.2024.17239a>
- Bowen, J. A., & Watson, C. E. (2024). *Teaching with AI: a practical guide to a new era of human learning*. Johns Hopkins University Press. <https://muse.jhu.edu/pub/1/monograph/book/123216>.
- Carson, J. G. & G. L. Nelson. (1996). Chinese students' perceptions of ESL peer response group interaction [J]. *Journal of Second Language Writing*, 5(1), 1-19.
- Stevenson, M., & Phakiti, A. (2014). The Effects of Computer-Generated Feedback on the Quality of Writing. *Assessing Writing*, 19, 51-65. <https://doi.org/10.1016/j.asw.2013.11.007>
- Gupta, A., Atef, Y., Mills, A., & Bali, M. (2024). Assistant, parrot, or colonizing loudspeaker? ChatGPT metaphors for developing critical AI literacies. *Open Praxis*, 16(1), 37-53. <https://doi.org/10.55982/openpraxis.16.1.631>
- Díaz, C. Ortiz, M., Gómez, J., & Sanhueza, S. (2023). English language assessment: Preservice teachers' self-reported views. *International Journal of Instruction*, 16(4), 405-422. <https://doi.org/10.29333/iji.2023.16424a>
- Ferris, D. R. & J. S. Hedgcock. (2014). *Teching L2 Composition: Purpose, Process, and Practice* (3rd edition) [M]. New York and London: Routledge/Taylor and Francis.
- Gao, Y., Y. Wang. & Christian, D.S. (2019). Research on the adoption of peer feedback comments on English writing and its influencing factors [J]. *Technology Enhanced Foreign Language Education*, (2), 17-24. <https://doi.org/10.20139/j.issn.1001-5795.2019.02.003>
- Hanjani A. M. (2016). Collaborative revision in L2 writing: learners' reflections [J]. *ELT Journal*, 296-307. <https://doi.org/10.1093/elt/ccv053>
- Huang, Y. S. (2006). A good essay is written: a discussion on English teaching [J]. *Foreign Language World*, 115(5), 13-16.
- Jacobs, G. M., A. Curtis., G. Braine. & S. Huang. (1998). Feedback on student writing: Taking the middle path [J]. *Journal of Second Language Writing*, 7(3), 307-317.
- Khonbi Z. A. & K. Sadeghi. (2012). The effect of assessment type (self vs. peer vs. teacher) on Iranian University EFL students' course achievement [J]. *Language Testing in Asia*, 4(2), 47-74.

- Lisa Sperber, Marit MacArthur, Sophia Minnillo, Nicholas Stillman & Carl Whithaus. (2025). Peer and AI Review + Reflection (PAIRR): A human-centered approach to formative assessment. *Computers and Composition*, 76, <https://doi.org/10.1016/j.compcom.2025.102921>.
- Liu, N. F. & D. Carless. (2006). Peer feedback: The learning element of peer assessment [J]. *Teaching in Higher Education*, 11(3), 279-290. <https://doi.org/10.1080/13562510600680582>
- Ma, X. M., J. H. Liu., & W. B. Du. (2023). Comparative study on the diagnostic effectiveness of English writing descriptive language scale and overall scoring scale based on peer assessment [J]. *Foreign Language Testing and Teaching*, (03), 46-55.
- Mangelsdorf, K. (1992). Peer reviews in the ESL composition classroom: What do the students think? [J]. *ELT Journal*, 46(3), 274-284.
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, Article 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- Price, M., Handley, K., & Millar, J. (2011). Feedback: Focusing attention on engagement. *Studies in Higher Education*, 36(8), 879–896. <https://doi.org/10.1080/03075079.2010.483513>
- Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & Olson, C. B. (2024). Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction*, 91. <https://doi.org/10.1016/j.learninstruc.2024.101894>
- Sun, S. G. (2019). Research on teacher-student cooperation evaluation based on output oriented method [D]. Beijing Foreign Studies University. <https://doi.org/10.26962/d.cnki.gbjwu.2019.000172>
- Torshizi, M. D. & M. Bahraman. (2019). I explain, therefore I learn: Improving students' assessment literacy and deep learning by teaching [J]. *Studies in Educational Evaluation*, 61: 66-73. <https://doi.org/10.1016/j.stueduc.2019.03.002>
- Tsui A. & M. Ng. (2000). Do secondary L2 writers benefit from peer comments? [J]. *Journal of Second Language Writing*, 9(2), 147-170.
- Wang, C., & Wang, Z. (2025). Investigating L2 writers' critical AI literacy in AI-assisted writing: An APSE model. *Journal of Second Language Writing*, 67, Article 101187. <https://doi.org/10.1016/j.jslw.2025.101187>
- Widiastuti, I. (2021). Assessment and feedback practices in the EFL classroom. [J] *Research and Evaluation in Education*, 7(1), 13-22. <https://doi.org/10.21831/reid.v7i1.37741>
- Wu, L. (2024). Application of generative artificial intelligence in middle school English writing teaching [J]. *Information Technology Education in Primary and Secondary Schools*, (08), 85-87.

Xu, Y. T. & G. T. L. Brown. (2016). Teacher assessment literacy in practice: A reconceptualization [J]. *Teaching and Teacher Education*, 58, 149-162. <http://dx.doi.org/10.1016/j.tate.2016.05.010>