Online-Journals.org

iJET International Journal of Emerging Technologies in Learning

iJET | eISSN: 1863-0383 | Vol. 18 No. 22 (2023) | OPEN ACCESS

https://doi.org/10.3991/ijet.v18i22.44699

PAPER

Impact of Peer Review on Learning Performance in a Smart Classroom Teaching Environment

Xue Liu(✉)

Foreign Languages and Tourism Department, Hebei Petroleum University of Technology, Chengde, China

cdpc_lx1@cdpc.edu.cn

ABSTRACT

With the rapid progress of educational technology, smart classrooms have gradually been widely applied, aiming to provide students with more efficient and innovative learning experiences. As a non-traditional assessment method, peer review has attracted widespread attention in this context, and its role in the learning process of students is increasingly prominent. However, there are still disputes and deficiencies regarding its specific applications and benefits in smart classroom environments. This study aimed to dig into the peer review standard and its score prediction in a smart classroom environment and evaluate the specific impact of peer review on learning performance. It is expected that this study can provide educators with a more accurate and practical peer review method, thereby optimizing the teaching and assessment modes of smart classrooms.

KEYWORDS

smart classroom, peer review, learning performance, score prediction, assessment standard

1 INTRODUCTION

With the continuous progress and digital development of educational technology, smart classrooms have gradually become an important component of the educational field, aiming to provide students with more innovative, efficient, and interesting learning experiences [1–8]. In such an environment, apart from the crucial role of teachers in providing teaching feedback to students, peer interaction and review also become increasingly important [9, 10]. Peer review, i.e., students assess each other's learning situations, provides valuable feedback for both parties, helping identify their deficiencies in learning and making timely adjustments [11, 12].

A smart classroom creates a more interactive, personalized, and efficient learning environment for teachers and students using modern technology and multimedia tools. In such an environment, the application of peer assessment has become a key teaching strategy, especially in English teaching. Students often need to participate in various interactions and discussions in the smart classroom. Peer assessment

Liu, X. (2023). Impact of Peer Review on Learning Performance in a Smart Classroom Teaching Environment. *International Journal of Emerging Technologies in Learning (iJET)*, 18(22), pp. 138–151. https://doi.org/10.3991/ijet.v18i22.44699

Article submitted 2023-07-13. Revision uploaded 2023-08-29. Final acceptance 2023-09-09.

encourages them to shift from passively receiving information to active assessment and reflection, which helps them deeply understand the language and cultural content. Meanwhile, students need to apply critical thinking skills to assess others' work. In English writing and speech assessment, the students need to analyze and assess the content, structure, grammar, and style, which exercises their analytical and critical abilities.

Peer review is not a novel concept in the educational field, but further research and exploration are needed to maximize its effectiveness in the context of smart classrooms [13–17]. Understanding and exploring the role of peer review in this specific environment not only helps educators design better teaching activities and assessment mechanisms, but also provides students with more targeted and effective learning feedback [18, 19]. In addition, peer review may also enhance students' learning motivation and improve their learning strategies and social skills [20].

However, existing studies of peer review mostly focus on its application, and there are still many disputes and deficiencies regarding its specific assessment standards and methods in smart classrooms [21, 22]. Several problems have not been solved, such as how to ensure the objectivity and fairness of peer review, how to accurately predict students' scores based on various indexes, and how to ensure the adaptability and flexibility of assessment methods [23–26].

This study focused on exploring the peer review standard and score prediction in a smart classroom teaching environment. A clear peer review standard was first defined, focusing on the speed at which peers answered questions and their mastery of knowledge points. The indexes of various test questions were used to predict scores, such as error and guess rate. Then this study evaluated the impact of peer review on learning performance in detail. After evaluating the individual learning performance of students, this study completed the similarity calculation of peer review and then proposed a targeted incentive strategy to improve the learning performance. This study aimed to provide educational practitioners with a more targeted peer review tool and method and promote better integration of teaching with assessment in the smart classroom, thereby improving the learning performance of students.

2 PEER REVIEW STANDARD AND SCORE PREDICTION IN A SMART CLASSROOM TEACHING ENVIRONMENT

A smart classroom emphasizes real-time interaction and feedback. In such a teaching environment, the response speed of students intuitively reflects their familiarity with relevant knowledge points and their thinking agility, thus providing a reference standard for other students. The mastery degree of knowledge points by peers is the core index for evaluating their learning effect. In a smart classroom, students deeply understand and master knowledge points through interactions and discussions. Therefore, the peers' mastery degree also indirectly reflects the effectiveness of teaching methods.

The peer review standard in a smart classroom teaching environment was determined by the speed of answering questions by peers and their mastery degree of knowledge points in this study. Such an assessment standard encourages students to participate in the class more actively and strengthens their interactions with other classmates, thereby promoting their learning depth and breadth. Compared with other possible assessment standards, the speed of answering questions and the mastery of knowledge points are more intuitive, easy to operate, and quantifiable, which not only makes peer review more concise and efficient and reduces subjective bias but also facilitates subsequent data analysis and research.

Let y_{MAX} be the maximum time for students to answer teachers' questions in a smart classroom teaching environment, and y_u be the time for the u-th student to answer questions, with $IS\_RI \in \{0,1\}$. For the time spent by students to answer teachers' questions, its weight y was calculated using the following equation:

$$y = (y_{MAX} - y_u) * IS\_RI \tag{1}$$

In practical situations, students' knowledge mastery is often a vague and continuous process. This study described the knowledge state as a fuzzy set to provide more detailed and accurate assessment. Let ϕ_u be the latent trait of the u-th student in the opinions of peers, s_{uj} be the differentiation degree of student u in knowledge point j in the opinions of peers, n_{uj} be the difficulty in answering knowledge point j by student u in the opinions of peers, F be the empirical scale constant in the opinions of peers, and s_{uj} be the mastery degree of knowledge point j by student u in the opinions of peers. The mastery degree of knowledge point j by student u was calculated based on the fuzzy cognitive diagnosis model (FCDM) using the following equation:

$$\beta_{uj} = \omega_j(u) = \frac{1}{1 + \exp[F * s_{uj}(\varphi_u - n_{uj})]} \tag{2}$$

Traditional cognitive diagnostic models use binarization (0 or 1) to represent whether students have mastered a certain knowledge point. However, FCDM uses the fuzzy set to describe students' knowledge state, which means that their mastery degree of a certain knowledge point can be a value between 0 and 1. The model defines a set of rules to infer students' knowledge states based on their performance on various test questions. These rules can be determined based on expert opinions or data-driven approaches. The model believes that students' knowledge mastery is a continuous process instead of just a simple dichotomy between mastery and non-mastery. Therefore, FCDM provides more detailed information on their knowledge mastery. Figure 1 shows the process of using the FCDM in this study.

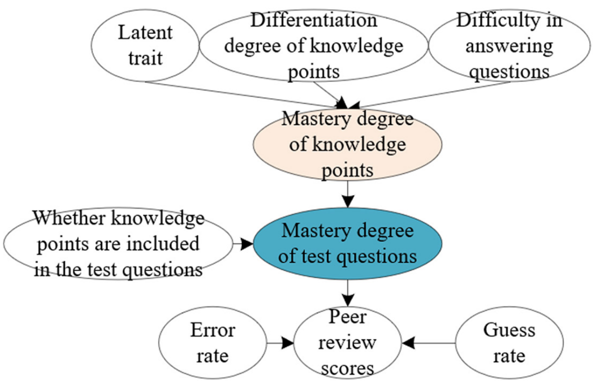

Fig. 1. FCDM process

Based on students' performance, a fuzzy set was defined for each knowledge point to describe their mastery degree of that knowledge point. According to the performance of students in various test questions and the diagnostic rules of the model, the mastery degree of each knowledge point was calculated. Let w_{uj} be whether u has examined the knowledge point j, and λ_{uj} be the mastery degree of test question k by u. The mastery degree of objective question k by u was considered as a fuzzy intersection of u's mastery degree of the knowledge points contained in k:

$$\lambda_{uk} = \bigcap_{k \le j \le J, w_{kj}=1} \omega_j(u) \tag{3}$$

The mastery degree of subjective question k by u was a fuzzy union of u's mastery degree of the knowledge points contained in k:

$$\lambda_{uk} = \bigcup_{k \le j \le J, w_{kj}=1} \omega_j(u) \tag{4}$$

Based on the output of the model, each student was provided with a description of their knowledge mastery, which was a fuzzy set or a specific numerical value. If the mastery degree of knowledge points by students was equivalent to the maximum and minimum values of u's mastery degree of all knowledge points examined by k, then there were:

$$i_{S\bigcap N}(z) = MIN(i_S(z), i_N(z)) \tag{5}$$

$$i_{S\bigcup n}(z) = MAX(i_S(z), i_N(z)) \tag{6}$$

Based on students' mastery degree of knowledge points, when predicting the score probability of peer review, various factors should be considered comprehensively, including error and guess rate when students answered test questions. Let $u(j)$ be u's mastery degree of j, $u(j)$ be u's mastery degree of all knowledge points involved in k, and w_{kj} be whether k includes j, then the prediction equation for subjective questions was as follows:

$$E_{uk} = \sum_{j=1}^{J} u(j) w_{kj} \tag{7}$$

Error rate is the probability that students give a wrong answer even though they have a definite grasp of knowledge points. The guess rate is the probability that students guess the correct answer without fully mastering the knowledge points. The error and guess rate of each student at each knowledge point were estimated by analyzing their answer data. Based on students' mastery degree of knowledge points and their error and guess rate, let a_k and h_k be the error and guess rate of test question u, respectively, E_{uk} be the scores of student u in test question k, λ_{uk} be the mastery degree of test question k by student u, and δ^2 be the variance obtained by normalizing the scores of subjective questions. The probability of predicted scores for objective and subjective questions was calculated as follows:

$$O\left(E_{uk} = 1 \middle| \lambda_{uk}, a_k, h_k\right) = (1 - a_k)\lambda_{uk} + h_k(1 - \lambda_{uk}) \tag{8}$$

$$O\left(E_{uk} \middle| \lambda_{uk}, a_k, h_k\right) = B\left(E_{uk} \middle| \left[\left(1 - a_k\right)\lambda_{uk} + h_k\left(1 - \lambda_{uk}\right)\right], \delta^2\right) \tag{9}$$

3 EVALUATING THE IMPACT OF PEER REVIEW ON LEARNING PERFORMANCE

3.1 Evaluating the individual learning performance

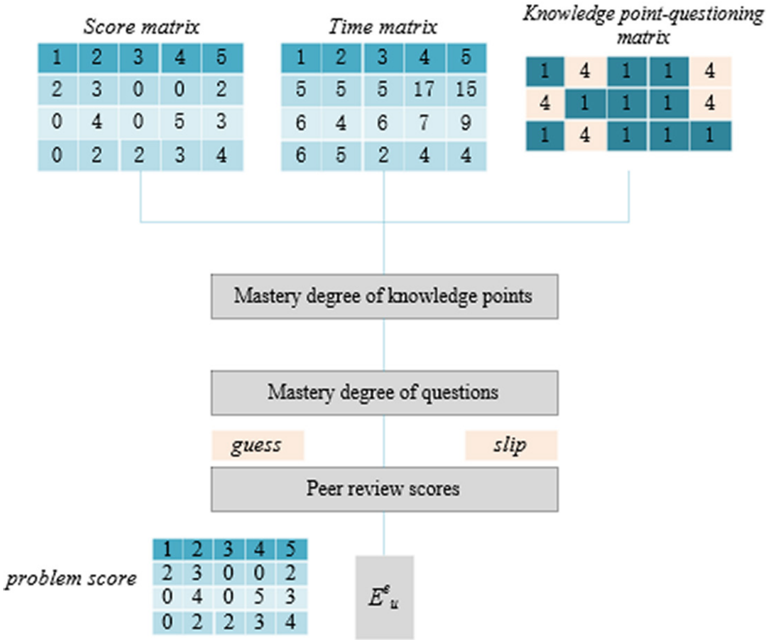

Fig. 2. Flowchart of obtaining peer review scores

The assessment of individual learning performance in a smart classroom teaching environment is a complex process that involves multiple assessment indexes and methods. Peer review scores reflect students' knowledge mastery and performance among their peers. Figure 2 shows the flowchart for obtaining peer review scores. Peer review reveals students' cooperation and communication abilities, as well as their contribution and knowledge mastery degree in group activities. Class score ranking is the relative position of students among all students in the entire class, which reflects their overall academic level and knowledge mastery. This study evaluated learning performance in two aspects: peer review scores and class score ranking. Let E_u^e be the scores of peer review text of student u to be evaluated, and l be the total number of students participating in peer review, then the peer review scores of students were calculated using the following equation:

$$E_u^e = \sum_{u=1}^{l} E_{uk} \tag{10}$$

Let E_u^z be the average ranking coefficient of student u, and l_u be the ranking of student u, then the class score ranking of students was calculated using the following equation:

$$E_u^z = 1 + \frac{l - l_u}{l - 1} \tag{11}$$

After normalizing the above results, the following calculation results were obtained:

$$\frac{E_u^z - E_{u\_MIN}^z}{E_{u\_MAX}^z - E_{u\_MIN}^z} \tag{12}$$

3.2 Calculating the peer review similarity

Peer assessment similarity plays an important role in the incentive strategy for improving learning performance, which mainly involves comparing the assessment results between students and exploring their consistency and differences in assessing others. If the peer assessment similarity is high, it means that students have a consistent assessment of a particular student or learning task, which enhances the fairness and reliability of peer assessment. If the similarity is low, there may be some biases or misunderstandings that require further in-depth analysis of their causes.

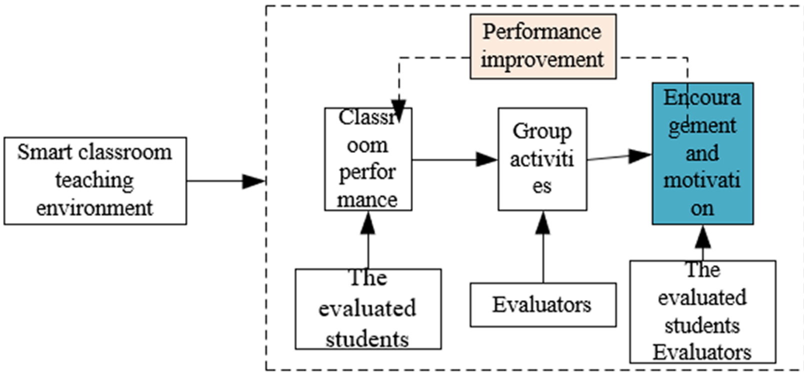

Fig. 3. Design of peer review tasks

Figure 3 shows the design process for peer assessment tasks. By comparing the assessment results of different students, it can be found that they share common standards and differences in the assessment, which helps teachers understand their assessment habits and standards. Understanding these common points and differences helps teachers provide clearer assessment guidance and suggestions for students. If a student obtains higher scores in the assessments of most peers but lower scores in the assessments of certain peers, this difference may require attention because it may indicate that there are some potential conflicts or misunderstandings that require mediation and guidance from teachers. If most students have similar assessments of a student, then teachers can be more confident in providing motivation or support to the student based on these assessments.

The specific steps for calculating peer review similarity were described in three aspects, namely, calculating the weight and semantic similarity of lexical items as well as the similarity between feature lexical item vectors. The long peer review text was first segmented into several meaningful lexical items. Let μ be the lexical item in the feature word, and $c_u = (\mu_{u1}, \mu_{u2}, ..., \mu_{ul})$ be the u-th peer review text. After collecting the segmented results, this study calculated the number of times that each lexical item occurred in the text, and the value of term frequency–inverse document frequency (TF-IDF). The obtained TF-IDF value was the weight of each lexical item. Let $QZ(\mu_u)$ be the weight of lexical items in peer review text, $yd_k(\mu_u)$ be the number of times that lexical item μ_u occurs in peer review text k, $yd(\mu_u)$ be the number of peer review texts where lexical item μ_u occurs, B be the total number of peer review texts, and $uyd(\mu_u)$ be the inverse document frequency of lexical item μ_u, then there was the following equation:

$$QZ(\mu_u) = yd(\mu_u) \times ufd(\mu_u) = yd_k(\mu_u) \times \log(B/fd(\mu_u)) \tag{13}$$

Lexical items were converted into word vectors using a pre-trained word vector model. For two lexical items, their word vectors were used to calculate cosine similarity, which gave the semantic similarity of both words. The value of cosine similarity is between –1 and 1. The closer it is to 1, the more similar it is. The closer it is to –1, the less similar it is. Let β be a regulation parameter, $DI(\mu_u, \mu_k)$ be the path length of lexical items in the semantic network, and $DI(\mu_u)$ be the depth of the node corresponding to lexical item μ_u in the semantic network from the root node, then there was the following equation:

$$SIM(\mu_u, \mu_k) = \frac{\beta \times (DE(\mu_u) + DE(\mu_k))}{(\beta \times (DE(\mu_u) + DE(\mu_k)) + DI(\mu_u, \mu_k))} \tag{14}$$

A feature vector was constructed for each assessment text based on the weight of the TF-IDF value. The dimension of the vector was the same as the number of lexical items, with the value of each dimension being the TF-IDF value of the corresponding lexical item. For two assessment texts, their feature vectors were used to calculate cosine similarity, which provided the overall similarity between both texts. For all assessment texts, a similarity matrix was calculated, with each value in the matrix representing the similarity between one text and another. Let $c_u = (\mu_{u1}, \mu_{u2}, \ldots, \mu_{ul})$, and $c_k = (\mu_{k1}, \mu_{k2}, \ldots, \mu_{kl})$; q be the weight value between two vectors c_u and c_k; and $VS(c_u, c_k)$ be the semantic similarity between two vectors, then there was the following equation:

$$TS(c_u, c_k) = g \times VS(c_u, c_k) + (1 - g)COSSIM(c_u, c_k) \tag{15}$$

The cosine similarity between two lexical items was calculated using the following equation:

$$COSSIM(c_u, c_k) = \frac{\sum_{v=1}^{\eta} QZ(\mu_{uv}) \times QZ(\mu_{kj})}{\sqrt{\sum_{v=1}^{l} \left(QZ(\mu_{uv})\right)^2 \times \sum_{f=1}^{b} \left(QZ(\mu_{kf})\right)^2}} \tag{16}$$

3.3 Incentive strategy for improving learning performance

In a smart classroom teaching environment, the formulation and implementation of a learning performance improvement incentive strategy are the key steps in the delicacy management of dynamic feedback during the learning process. In this context, it is particularly important to provide a personalized incentive strategy for students based on a comprehensive evaluation of peer review similarity and score ranking. The implementation plan for the incentive strategy is described in detail below.

Step 1: Data collection and processing. Advanced natural language processing technology was used to deeply analyze students' peer review texts, extract key information, and construct feature lexical item vectors. After obtaining students' learning performance data, such as exam scores, homework completion status, etc., their score ranking was generated based on this.

Step 2: Peer assessment similarity analysis. By calculating the semantic similarity between the assessment content of each student and the assessments of other students, a similarity matrix was formed to identify assessment patterns and tendencies. Combined with TF-IDF and the pre-trained word vector model, accurate semantic similarity was calculated.

Step 3: Comprehensive evaluation and classification of students. Based on peer review similarity and score ranking, a clustering analysis method was used to classify students into different types, with each type representing a specific learning mode and performance feature. A machine learning algorithm was used to classify students automatically, ensuring the accuracy and fairness of the classification. The feature similarity of students was calculated using the following equations:

$$L_{uk} = \frac{1}{b}\sum_{j=1}^{b} SIM(E_u^j, E_k^j) \tag{17}$$

$$SIM(E_u^j, E_k^j) = COS(E_u^j, E_k^j) = \frac{E_u^j \cdot E_k^j}{\left\|E_u^j\right\| \cdot \left\|E_k^j\right\|} \tag{18}$$

Step 4: Formulation and implementation of the incentive strategy. A specific incentive strategy was formulated for each type of student based on their classification results. The strategy should consider the learning styles, needs, strengths, and weaknesses of the students and may include additional resource support, tutoring plans, learning partner matching, and guidance on learning methods. The effect of the incentive strategy was regularly evaluated to ensure its effectiveness, and the strategy was fine-tuned and optimized based on feedback.

Step 5: Continuous monitoring and feedback. During the implementation process of the incentive strategy, students' feedback and learning data were continuously collected to ensure that the strategy matched their actual needs. Students' classification and the incentive strategy were regularly updated using the data-driven method to ensure the timeliness and adaptability of the strategy.

4 EXPERIMENTAL RESULTS AND ANALYSIS

Figure 4 shows the proximity between different learning performance improvement incentive strategies and teachers' scores under different numbers of peer reviews. As the number of peer reviews increases, the proximity between all strategies and teachers' scores increases, which indicates that the increase in peer review times improves the scoring accuracy. Starting from the 50th peer review, the proximity of the proposed strategy in this study reaches 92.3%, significantly higher than that of the other two strategies. In the 380th peer review, the proximity reaches 97.0%, which is also the highest among the three strategies. Therefore, the increase in peer review times improves the proximity between peer review scores and teachers' scores, which is applicable to all strategies. Although the proximity between all strategies and teachers' scores shows an increasing trend as the number of peer reviews increases, the effect of different strategies is significantly different. In summary, the proposed learning performance improvement incentive strategy shows high proximity to teachers' scores in peer review, proving the effectiveness and practicability of this strategy.

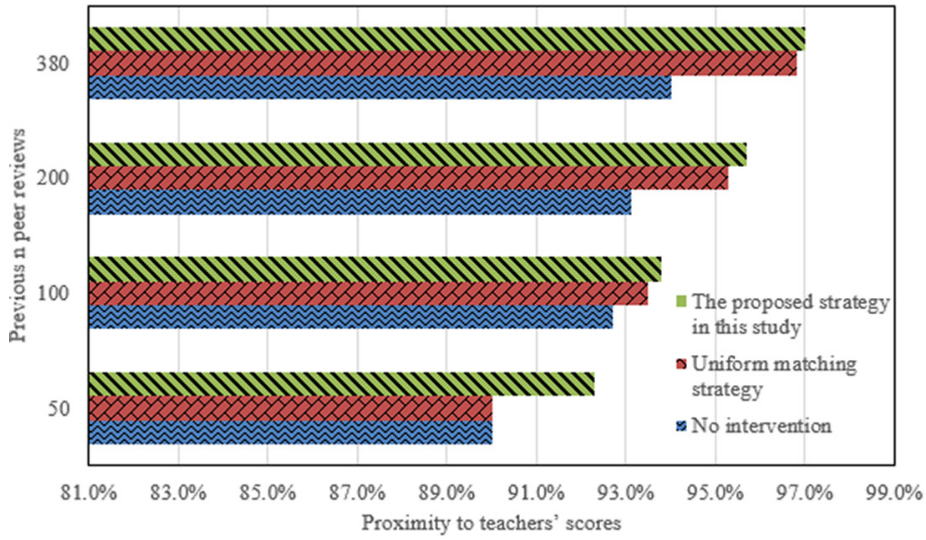

Fig. 4. Proximity between peer review and teachers' scores under different learning performance improvement incentive strategies

Table 1 shows the root mean square error (RMSE) values of peer review scores and true scores. It can be seen from the table that the proposed strategy has the lowest RMSE value in most cases overall, indicating relatively high prediction accuracy. The RMSE value of the uniform matching strategy lies between that of the proposed strategy and the no intervention strategy in most cases, indicating moderate prediction accuracy. The no intervention strategy often has the highest RMSE value, which means its prediction accuracy is relatively low. As the number of peer reviews increases, there is no obvious overall trend, such as a stable increase or decrease in RMSE values. However, some local fluctuations can be observed because each peer review is affected by the difficulty of test questions, the participation of students, and other factors not being considered.

Table 1. RMSE values of peer review scores and true scores

The n-th Peer Review Incentive Strategies	The Proposed Strategy in this Study	No Intervention	Uniform Matching Strategy
1	2.23	3.12	2.91
2	3.25	4.38	3.51
3	1.69	2.21	1.88
4	3.18	3.64	3.34
5	2.16	2.73	2.41
6	2.32	2.74	2.65
7	3.59	3.62	3.71
8	4.05	4.20	4.31
9	1.62	1.93	1.66
10	2.33	2.54	2.48

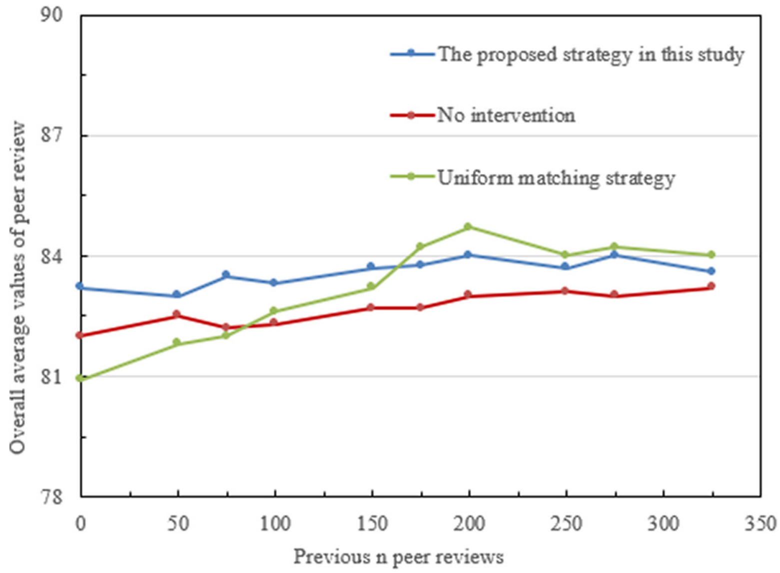

Fig. 5. Overall average values of peer review using different learning performance improvement incentive strategies

Based on Figure 5, the overall average values of peer review using different learning performance improvement incentive strategies can be analyzed. It can be seen from the figure that the average values of peer review using all strategies increase as the number of peer reviews increases, meaning that students' performance in peer review gradually improves over time and with the accumulation of experience, which reflects better learning performance. Although the average values of peer review using all strategies increase as the number of peer reviews increases, there are differences in their upward trends and volatility. The proposed strategy maintains a stable and high average value of peer review on the whole because it combines multiple assessment standards and performance incentive strategies. Although the no intervention strategy starts from a lower starting point, it shows a stable upward trend throughout the entire process, indicating that students' performance in peer review gradually improves even without specific incentive measures. The uniform matching strategy shows a clear growth trend as the number of peer reviews increases, indicating that uniform matching of students and assessment of them according to specific standards help improve their learning performance.

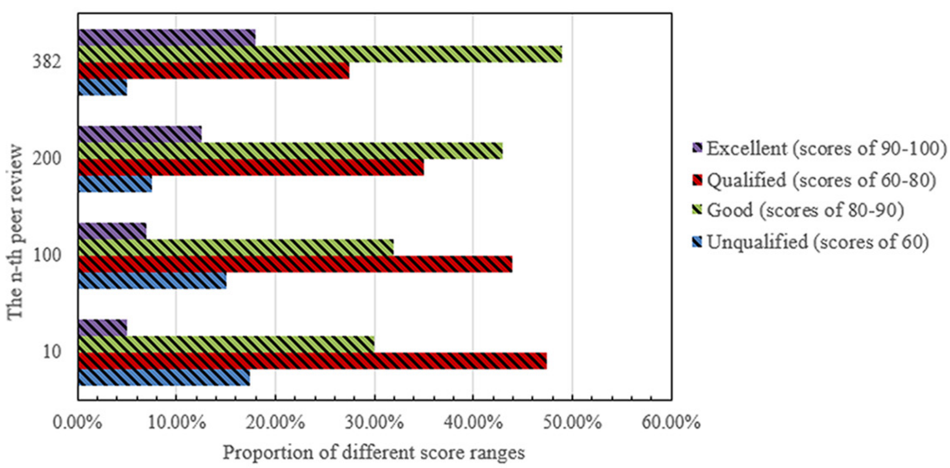

Fig. 6. Number of peer reviews and proportion of different score ranges

Table 2. Learning performance of the experimental class at different incentive stages

Dimensions	Test Names	Average Value	Standard Deviation	t-Value	p-Value
Learning performance in the pre-incentive stage	Pretest	4.88	0.26	–6.628	0.000**
	Posttest	5.26	0.15		
Learning performance during the incentive stage	Pretest	4.95	0.25	–7.824	0.000**
	Posttest	5.31	0.29		
Learning performance in the post-incentive stage	Pretest	4.99	0.18	–6.315	0.003**
	Posttest	5.63	0.23		

Based on Figure 6, the number of peer reviews and the proportion of different score ranges can be analyzed. As the number of peer reviews increases, the proportion of unqualified students significantly decreases from 17.5% to 5.0%, which indicates that peer review significantly improves learning performance. Similarly, the proportion of qualified students also decreases from 47.5% to 27.5%, which indicates that students have changed from being unqualified to qualified, and even a considerable number of them have further improved their performance and entered a higher score range. As the number of peer reviews increases, the proportion of students in the good score range significantly increases from 30.0% to 49.0%, which means that more students have improved their learning performance through peer review and achieved good scores. As the number of peer reviews increases, the proportion of students in the excellent score range also significantly increases from 5.0% to 18.0%, which indicates that the peer review incentive strategy is very effective in stimulating students' potential and driving them to achieve optimal learning performance.

Based on Table 2, the learning performance values of the experimental class at different incentive stages can be analyzed. In the pre-incentive stage, the average value of learning performance increases from 4.88 to 5.26, indicating that learning performance significantly improves in the experiment during the pre-incentive stage. The t-value is –6.628, and the p-value is less than 0.01, indicating that the learning performance difference before and after the experiment is significant during the pre-incentive stage. During the incentive stage, the average value of learning performance increases from 4.95 to 5.31, indicating a significant improvement in learning performance in the experiment during the incentive stage. The t-value is –7.824, and the p-value is less than 0.01, indicating a significant difference in learning performance before and after the experiment during the incentive stage. In the post-incentive stage, the average value of learning performance increases from 4.99 to 5.63, with the most significant improvement among the three stages, indicating the important impact of incentive on learning performance. The t-value is –6.315, and the p-value is less than 0.01, indicating a significant difference in learning performance before and after the experiment in the post-incentive stage. The distribution of experimental data is relatively stable in terms of standard deviation. The standard deviation slightly increases in the learning performance experiment in the post-incentive stage, indicating a more dispersed performance distribution among students. The above analysis conclusions indicate that the incentive strategy has a positive impact on students' learning performance in all stages, and its effect is the most prominent, especially in the post-incentive stage.

5 CONCLUSION

This study evaluated the impact of peer review on learning performance in detail. After evaluating the individual learning performance of students, this study completed the similarity calculation of peer review and then proposed a targeted incentive strategy to improve the learning performance. It can be seen from the data analysis that the strategy proposed in this study outperforms other strategies in terms of proximity between peer review and teachers' scores, RMSE value, and the overall average value of peer review. As the number of peer reviews increased, students' learning performance improved, indicating the positive role of peer review in improving learning performance in a smart classroom environment. The experimental results of the incentive strategy showed that students' learning performance significantly improved in different incentive stages, with the most significant improvement especially in the post-incentive stage.

Based on the research findings of this study, it can be concluded that peer review not only serves as an effective assessment tool in a smart classroom teaching environment but also significantly improves students' learning performance. The peer review strategy proposed in this study performs well in predicting peer review scores, calculating peer review similarity, and the learning performance improvement incentive strategy, and has advantages over other strategies. In addition, as peer review continues, the learning performance of students continues to improve, confirming the long-term positive impact of peer review. Therefore, peer review in a smart classroom environment is worth further promotion and application.

6 REFERENCES

[1] N. Songkram, N. Songkram, S. Chootongchai, and T. Samanakupt, "Developing students' learning and innovation skills using the virtual smart classroom," *International Journal of Emerging Technologies in Learning*, vol. 16, no. 4, pp. 34–51, 2021. https://doi.org/10.3991/ijet.v16i04.15221

[2] M. N. Y. Utomo, M. Sudaryanto, and K. Saddhono, "Tools and strategy for distance learning to respond COVID-19 pandemic in Indonesia," *Ingénierie des Systèmes d'Information*, vol. 25, no. 3, pp. 383–390, 2020. https://doi.org/10.18280/isi.250314

[3] S. Jacques and T. Lequeu, "The attractiveness of reversing teaching forms—feedback on an electrical engineering course," *International Journal of Engineering Pedagogy*, vol. 10, no. 3, pp. 21–34, 2020. https://doi.org/10.3991/ijep.v10i3.12361

[4] H. Su, "A research-based smart classroom in an exercise physiology course," *International Journal of Emerging Technologies in Learning*, vol. 16, no. 18, pp. 97–112, 2021. https://doi.org/10.3991/ijet.v16i18.25339

[5] L. M. Liu, "Analysis on class participation based on artificial intelligence," *Revue d'Intelligence Artificielle*, vol. 34, no. 3, pp. 369–375, 2020. https://doi.org/10.18280/ria.340316

[6] S. Tsoukos, P. Lazos, P. Tzamalis, A. Kateris, and A. Velentzas, "How effectively can students' personal smartphones be used as tools in physics labs?" *International Journal of Interactive Mobile Technologies*, vol. 15, no. 14, pp. 55–71, 2021. https://doi.org/10.3991/ijim.v15i14.22375

[7] M. Lakshaga Jyothi and R. S. Shanmugasundaram, "Design and implementation of intelligent classroom framework through light-weight neural networks based on multimodal sensor data fusion approach," *Revue d'Intelligence Artificielle*, vol. 35, no. 4, pp. 291–300, 2021. https://doi.org/10.18280/ria.350403

[8] P. J., A. L. R., and A. C. M. T, "Accelerating the move towards online learning through cloud platforms in higher education sectors using smart devices during COVID-19," *International Journal of Interactive Mobile Technologies*, vol. 15, no. 10, pp. 33–48, 2021. https://doi.org/10.3991/ijim.v15i10.22163

[9] W. Groeneveld, J. Vennekens, and K. Aerts, "Engaging software engineering students in grading: The effects of peer assessment on self-evaluation, motivation, and study time," *arXiv Preprint*, 2020. https://doi.org/10.48550/arXiv.2012.03521

[10] I. Horikoshi and Y. Tamura, "Analysis of 'evaluation behavior' using students' peer assessment process data," in *ICCE 2019-27th International Conference on Computers in Education, Proceedings*, vol. 2, 2019, pp. 761–764.

[11] N. Kiron and J. Vassileva, "Tower of questions: Gamified testing to engage students in peer review," in *CEUR Workshop Proceedings*, vol. 2354, 2019, pp. 113–119.

[12] K. Willey, "Self and peer review to engage students with assessment standard, develop their judgement and critical evaluation skills and extend the benefits of instructor feedback," in *Proceedings of the 46th SEFI Annual Conference 2018: Creativity, Innovation and Entrepreneurship for Engineering Education Excellence*, 2019, pp. 512–519. https://doi.org/10.21125/iceri.2017.2023

[13] S. Jie and J. Cross, "Peer evaluation of academic presentations by Japanese graduate students: Qualitative evaluation," in *2015 International Conference on Advanced Mechatronic Systems (ICAMechS)*, Beijing, China, 2015, pp. 174–178. https://doi.org/10.1109/ICAMechS.2015.7287104

[14] S. Kulturel-Konak, A. Konak, G. E. O. Kremer, I. Esparragoza, and G. Yoder, "Peer review and assessment resource (PEAR) to assess students' professional skills," in *IIE Annual Conference and Expo*, 2014, pp. 746–753.

[15] A. Settle, C. Wilcox, and C. Settle, "Engaging game design students using peer review," in *SIGITE'11-Proceedings of the 2011 ACM Special Interest Group for Information Technology Education Conference*, 2011, pp. 73–78. https://doi.org/10.1145/2047594.2047614

[16] J. K. L. Poon, "Students' perceptions of peer review in project work," *Conferences in Research and Practice in Information Technology Series*, vol. 114, pp. 87–93, 2011.

[17] G. Nanda, S. Wei, A. Katz, C. Brinton, and M. Ohland, "Work-in-progress: Using latent Dirichlet allocation to uncover themes in student comments from peer evaluations of teamwork," in *2022 ASEE Annual Conference & Exposition*, 2022. https://peer.asee.org/41066.pdf

[18] A. Mitchell, M. Scott, J. Walton-Rivers, M. Watkins, W. New, and D. Brown, "An exploratory analysis of student experiences with peer evaluation in group game development projects," in *Proceedings of the 2022 Conference on United Kingdom & Ireland Computing Education Research*, New York, USA, pp. 1–7, 2022. https://doi.org/10.1145/3555009.3555021

[19] M. E. Morales-Trujillo, M. Galster, F. Gilson, and M. Mathews, "A three-year study on peer evaluation in a software engineering project course," *IEEE Transactions on Education*, vol. 65, no. 3, pp. 409–418, 2021. https://doi.org/10.1109/TE.2021.3123682

[20] J. Wang and P. K. Imbrie, "Students' peer evaluation calibration through the administration of vignettes," in *2010 Annual Conference & Exposition*, Louisville, Kentucky, pp. 15–1136, 2010. https://doi.org/10.18260/1-2--16755

[21] S. Ghada and A. Beena, "Using peer review to assess individual performance in team projects for freshman engineering students in the middle east," in *ASEE Annual Conference and Exposition, Conference Proceedings, 118th ASEE Annual Conference and Exposition*, 2011.

[22] S. Ghada and A. Beena, "Using peer review to assess individual performance in team projects for freshman engineering students in the middle east," in *SEFI Annual Conference*, pp. 890–898, 2011.

[23] Y. Peng, Y. Li, Y. Su, K. Chen, and S. Jiang, "Effects of group awareness tools on students' engagement, performance, and perceptions in online collaborative writing: Intergroup information matters," *The Internet and Higher Education*, vol. 53, p. 100845, 2022. https://doi.org/10.1016/j.iheduc.2022.100845

[24] C. H. Zhou, S. Q. Wei, and M. W. Ohland, "Team tenure—the longitudinal study of engineering student peer rating quality," in *ASEE Annual Conference and Exposition, Conference Proceedings, 2022 ASEE Annual Conference and Exposition: Excellence through Diversity*, 2022.

[25] K. Misiejuk and B. Wasson, "Backward evaluation in peer assessment: A scoping review," *Computers & Education*, vol. 175, p. 104319, 2021. https://doi.org/10.1016/j.compedu.2021.104319

[26] L. Huang, "From feedback to revision–how can collaborative peer feedback make a difference?" in *2021 Tenth International Conference of Educational Innovation through Technology (EITT)*, Chongqing, China, pp. 228–232, 2021. https://doi.org/10.1109/EITT53287.2021.00052

7 AUTHOR

Xue Liu is an accomplished scholar who completed her Bachelor of Arts degree at the School of Foreign Languages, Binzhou University, Binzhou, China, in 2008. She further broadened her academic horizons by earning her Master's degree from the School of Foreign Studies at China University of Petroleum, located in Qingdao, China, in 2010. Currently, she is a respected member of the faculty at the Foreign Languages and Tourism Department, Hebei Petroleum University of Technology. Her research interests span a wide range of topics including English teaching evaluation methods, effective English teaching strategies, and linguistics. Her contribution to these research areas are both significant and impactful (E-mail: cdpc_lx1@cdpc.edu.cn; ORCID: https://orcid.org/0000-0002-8327-8336).