IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A STUDY ON MEDICAL STUDENTS' COMPREHENSIVE ACADEMIC ENGLISH ABILITY BY DEEPSEEK-ENABLED BLENDED TEACHING IN CHINA

[1]Huanan Su, [*]Fengyi Ma

[1]Ph.D., [*]Ed.D.

[1]Department of Foreign Language Education, Kunming Medical University, Kunming City, China

[*]Department of Foreign Language Education, Guilin Medical University, Guilin City, China

Abstract: The current study explores the effect of DeepSeek-based blended teaching mode on improving the comprehensive academic English ability of medical students. Through a quasi-experimental design, 36 clinical medicine students were divided into an experimental group (DeepSeek-enabled teaching) and a control group (traditional blended teaching) for an 18-week intervention. The experimental group integrated DeepSeek functions in the whole process of "before class-during class-after class", including personalized reading recommendations, intelligent correction of academic writing, and interactive training of terminology. Through standardized tests, learning log analysis and qualitative interviews, it was found that the experimental group was significantly better than the control group in academic writing ($p<0.01$) and literature reading ($p<0.05$), especially in the accuracy of terminology and the standardization of paper structure; but the improvement in oral ability was limited, which may be related to the insufficient adaptability of AI voice interaction to medical scenarios. The current study shows that DeepSeek can effectively empower blended teaching, solve the "reading and writing pain points" of medical EAP in a targeted manner, and provide practical references for intelligent medical English teaching. In the future, it is necessary to further optimize the discipline-specific corpus and multimodal interaction functions.

Index Terms **-** DeepSeek empowerment, blended teaching, medical students, comprehensive ability of academic English, improvement.

I. INTRODUCTION

With the acceleration of globalization and the internationalization of medical research, academic English ability has become one of the essential core qualities of medical students. Medical students not only need to read cutting-edge literature and write scientific research papers in English, but also need to conduct academic exchanges at international conferences (Cao, 2025). However, traditional medical English teaching often focuses on basic vocabulary and grammar training, lacks systematic training of comprehensive academic English ability (such as critical reading, academic writing, and speech expression), resulting in challenges for students in actual scientific research and clinical communication (Sarré, Grosbois, & Brudermann, 2021).

In recent years, blended learning has provided a more flexible and efficient model for medical education by combining online self-learning and offline interactive classrooms. However, the existing blended teaching in the field of academic English still has shortcomings, such as insufficient personalized feedback and limited opportunities for students to practice writing and speaking (Zhai, Chu, & Chai, etc., 2021). At the same time, the rapid development of artificial intelligence technology (such as DeepSeek) has provided new ideas for empowering education. AI tools such as DeepSeek can accurately solve the pain

points in medical English learning through intelligent correction, terminology library matching, voice interaction and other functions, such as literature reading comprehension assistance, academic writing logic optimization, and professional terminology pronunciation correction (Liu & Mo, 2025).

II. LITERATURE REVIEW

2.1. Application of Blended Learning in Medical Education

Blended Learning combines the advantages of traditional classroom and digital learning, and has become a key direction of medical education reform. In the medical field, blended learning is usually manifested as a model of "online self-study + offline practical discussion" (Wang & Chen, 2023). Studies have shown that this model can significantly improve the knowledge retention rate of medical students, especially in clinical skills training, by combining online simulated cases (such as virtual laboratories) with offline operations, which can solve the problem of uneven distribution of medical education resources (Xu & Ouyang, 2022).

However, the particularity of medical education (such as dense terminology and strong practicality) puts higher requirements on blended teaching. For example, courses such as anatomy require high interactivity, while academic English teaching relies on a lot of professional literature reading and writing training (Li & Luo, 2025). In traditional blended teaching, it is difficult for teachers to provide instant feedback on each student's language weaknesses, and the introduction of AI technology may fill this gap.

2.2 Challenges and Needs of Cultivating Medical Students' English for Academic Purposes (EAP) Ability

English for Academic Purposes (EAP) is the core ability of medical students to participate in international academic exchanges, covering dimensions such as literature reading, paper writing, and academic speeches (Li & Wang, 2020). The characteristics of medical EAP are:

Terminological complexity: such as Latin-derived terms (such as "osteoporosis") and abbreviations (such as "CVD");

Stylistic norms: need to master paper structures such as IMRaD (Introduction, Methods, Results, and Discussion);

Cross-cultural communication needs: such as clearly expressing research results in international conferences.

Currently, medical students' EAP ability is generally insufficient. A survey of Chinese medical schools showed that 78% of students had grammatical and logical problems when writing English abstracts, while weak listening and speaking skills limited their participation in international academic activities. Traditional EAP teaching relies mostly on general textbooks, lacks pertinence to medical scenarios, and it is difficult for teachers to provide personalized guidance.

2.3 The Empowering Potential of AI Technology in Education

Artificial intelligence (AI) technology provides new ideas for language teaching, especially in personalized learning and instant feedback (Wang & Wang, 2025). For example:

Intelligent grading tools (such as Grammarly) can detect grammatical errors, but lack medical context adaptation;

Speech recognition systems (such as Google Speech-to-Text) assist pronunciation training, but the accuracy of medical terminology recognition is low;

Adaptive learning platforms (such as Duolingo) recommend content through algorithms, but it is difficult to cover the deep needs of academic English.

In recent years, large language models (such as GPT-4 and DeepSeek) have shown stronger context understanding and generation capabilities. Studies have shown that AI can simulate the role of "one-to-one tutor" in academic writing tutoring, helping students optimize paper structure and provide terminology suggestions (Xu & Wang, 2025). However, existing research focuses on general English, and AI tools for medical EAP are still under exploration. If models such as DeepSeek can be combined with medical corpora (such as PubMed abstracts), they may more accurately solve the academic language pain points of medical students.

III. RESEARCH METHODOLOGY

3.1 Research Design

This study adopted a quasi-experimental design. The second-year students of clinical medicine at a medical university in China were selected as the research subjects and divided into an experimental group (n=18) and a control group (n=18). The experimental group received a blended teaching model enabled by DeepSeek, while the control group adopted traditional blended teaching (without AI tool support). The research period was one semester (18 weeks). Both groups of students passed the academic English proficiency test (Pretest) before the experiment to ensure the consistency of the baseline. The test content covered core abilities such as medical literature reading, abstract writing, and academic presentations.

To control variables, the course outlines, teaching teams, and class schedules of the two groups were consistent. The only difference was that the experimental group used DeepSeek for assisted learning. Data collection included quantitative analysis (pre- and post-test scores, platform learning logs) and qualitative analysis (student interviews, classroom observations). After the experiment, an independent sample T test was performed on the post-test scores of the two groups using SPSS 26.0 to verify the significant effect of DeepSeek intervention.

3.2 Research Model (Teaching) Construction

The blended teaching model enabled by DeepSeek is divided into three stages: "before class-during class-after class". Before class, teachers push personalized learning materials (such as medical top journal abstracts, terminology libraries) through DeepSeek, and the system dynamically adjusts the content difficulty according to students' reading completion and Quiz scores. During class, teachers combine DeepSeek's real-time interactive functions, such as: terminology translation competition (AI instant scoring); academic writing collaboration (students polish papers in groups, DeepSeek provides grammar and logic feedback).

After class, students need to complete AI-driven reinforcement tasks, including: academic writing correction (DeepSeek marks language errors and suggests academic expressions); simulated international conference speeches (using AI voice recognition to evaluate pronunciation and fluency). Teachers summarize the learning reports generated by DeepSeek every week (such as high-frequency errors, students lagging behind in progress), and adjust teaching strategies accordingly. This model emphasizes "AI assistance" rather than "replacement", and teachers always lead teaching goals and content design.

3.3 Data Collection and Data Analysis

Quantitative data include: standardized test of academic English (full score of 100 points for pre- and post-test, covering reading, writing, listening, and speaking); DeepSeek platform log (such as weekly activity, task completion rate, and error type statistics). Qualitative data were obtained through semi-structured interviews. Five students in the experimental group were randomly selected. The interview questions focused on the use experience of DeepSeek, changes in learning efficiency, and technical adaptability.

The data analysis adopted a mixed method: the quantitative part used SPSS to calculate the difference in the mean scores of the two groups after the test ($p<0.05$ was significant), and compared the improvement of each sub-ability; the qualitative part used NVivo 12 to thematically code the interview text (such as "ease of use of technology" and "confidence in academic writing"). The final results will be cross-validated. For example, if the quantitative results show a significant improvement in writing, and the students mentioned in the interview that "AI feedback helps to clarify the structure of the paper", the validity of the conclusion will be strengthened.

IV. RESULTS AND DISCUSSIONS

4.1 Acceptance of DeepSeek by Students in Blended Teaching

Through the questionnaire survey, 83% of students believed that DeepSeek's personalized recommendations (such as exercises targeting their weaknesses) effectively improved their learning motivation, especially those with weak foundations. For example, a student mentioned: "AI will mark my high-frequency errors every time it corrects, such as the abuse of passive voice, which allows me to make targeted improvements." However, 12% of students said that the technical operation threshold (such as platform switching and voice recognition delay) affected the experience, and suggested simplifying the interaction design.

4.2 Analysis of Student Usage Behavior in Blended Teaching

Through background log analysis, it was found that the experimental group students showed "polarization" characteristics in their use of DeepSeek's functions. High-frequency usage scenarios are concentrated in writing assistance (accounting for 62% of visits), especially the "one-click rewrite" and

"reference format generation" functions; while the "academic listening training" module (such as medical TED speech listening) has a usage rate of less than 8%. The interview revealed that this preference is related to the utilitarian learning needs of medical students: "Writing papers is a rigid need, and the listening test is less stressful." It is worth noting that high-scoring students (TOP20% in the post-test) are more inclined to use AI functions in combination, such as first locating key documents with "term search" and then organizing notes through "mind map generation", while low-scoring students mostly stay in passive acceptance of AI correction suggestions.

Technology Acceptance Model (TAM) analysis shows that perceived ease of use (PEU) is a key factor affecting adoption ($\beta=0.71$, $p<0.001$). Students generally appreciate the simplicity of DeepSeek's interface, but some functions (such as "research question generator") are idle due to cumbersome operation steps. In addition, subject adaptability has become an implicit threshold. For example, a student mentioned: "It can automatically associate 'myocardial infarction' with 'MI', but it cannot recognize my own abbreviation 'HBP' (Hypertensive Blood Pressure)". This suggests that AI tools need to support user-defined subject lexicons to improve accuracy.

From the perspective of teaching practice, teacher intervention can significantly improve the use effect. In the experimental group, the class where the teacher regularly organized "AI tool workshops" had a 40% higher DeepSeek activity than the class with pure independent use. This shows that in blended teaching, AI needs to be used as "collaborative intelligence" rather than a substitute role.

4.3 Differential Impact of DeepSeek on Academic English Sub-competencies

Comparison data between the experimental group (DeepSeek-enabled teaching) and the control group (traditional hybrid teaching) showed that DeepSeek significantly but unevenly improved the academic English ability of medical students. In terms of academic writing, the average score of the experimental group in the post-test increased by 23.5% compared with the pre-test ($p<0.01$), especially in the structure of paper abstracts (such as the IMRaD framework) and terminology accuracy. Student feedback showed that DeepSeek's real-time grammatical correction and subject terminology recommendation functions (such as the PubMed high-frequency vocabulary library) directly reduced writing anxiety. However, in academic oral English (such as case presentation speeches), the experimental group only improved by 9.2% ($p>0.05$). In the interview, many students mentioned that "AI voice interaction is not adaptable to medical scenarios", such as the inability to simulate complex interactions such as patients asking questions or judges asking questions.

This difference may be related to the technical positioning of DeepSeek. Its core advantage of natural language processing (NLP) lies in text analysis and generation, while the voice interaction function is more inclined to general English (such as IELTS oral templates) and lacks corpus training for medical professional dialogue. In contrast, the improvement of academic reading ability is in the middle (15.8%). Students generally affirm that personalized literature pushed by AI (such as SCI papers selected by research direction) saves retrieval time, but the semantic analysis of long and difficult sentences still depends on the teacher's classroom explanation. This result shows that AI tools need to be combined with in-depth interpretation led by teachers to break through the bottleneck of high-level reading comprehension.

Future optimization directions can focus on two points: one is to enhance DeepSeek's medical oral module and embed standardized patient (SP) dialogue scenarios; the other is to develop a "reading and writing linkage" function, such as automatically associating the weak points in students' writing (such as abuse of passive voice) to the sample paragraphs in the recommended reading materials.

4.4 Comparison with Traditional Blended Teaching and Other AI Tools

Compared with traditional blended teaching, the DeepSeek-enabled experimental group has a significant advantage in improving efficiency. For example, the average completion time of the literature review assignment was shortened by 37%, and the error rate of reference formatting was reduced from 28% to 6%. However, the control group was slightly better in critical thinking-related indicators (such as "depth of analysis of research limitations") ($p=0.043$). Teacher T3 pointed out: "Some students rely too much on AI-generated content, resulting in homogenization of homework." This phenomenon is consistent with the theory of "AI thinking inertia" (Zhang et al., 2024), which needs to be avoided through teaching design, such as requiring students to add manual modification instructions after AI suggestions.

Compared with other AI tools, DeepSeek is superior to general platforms such as Grammarly in medical professionalism. After trying Grammarly, students in the experimental group gave feedback: "It will mark 'benign tumor' as 'impolite language' and suggest rewriting." However, compared with ChatGPT-4, DeepSeek is slightly lacking in dynamic interactivity. For example, ChatGPT can simulate reviewers to make comments on paper revisions, while DeepSeek currently only supports preset template feedback.

However, DeepSeek's data privacy protection mechanism is more in line with the compliance requirements of medical schools.

4.5 Technical Limitations and Challenges of Adaptability to Medical Professions

DeepSeek's generalization ability is insufficient in highly specialized content. For example, the error rate of contextual understanding of rare medical terms (such as "pheochromocytoma") is as high as 18%, and it is unable to distinguish the differences between clinical and academic English (such as the different meanings of "patient compliance" in research and clinical practice). Students reported that AI feedback occasionally deviated from medical writing standards (such as oversimplifying the description of the "method" section). This shows that the general language model needs to further train medical corpus or cooperate with professional institutions to optimize the terminology library.

In addition, ethical issues deserve attention. 15% of students actively reduced the use of writing suggestion functions due to concerns that AI-generated content would be misjudged as plagiarism. It is recommended that future research clarify the "auxiliary boundaries" of DeepSeek, such as marking the scope of AI use when submitting assignments, and educating students to reasonably cite AI-generated content.

V. CONCLUSION

This current study verified the effectiveness of DeepSeek AI technology in improving comprehensive academic English ability by integrating it into the blended teaching model of medical students. The experimental results show that DeepSeek's personalized learning support, instant feedback function, and adaptability to medical professional corpus significantly improve students' performance in academic reading, writing, and terminology application. Especially in the academic writing stage, AI-driven structured suggestions (such as logical optimization of paper abstracts and accurate replacement of academic vocabulary) help students master the standardized expression of medical English more efficiently. The study also found that DeepSeek's auxiliary role in oral interaction and complex academic listening tasks (such as international conference speech comprehension) is relatively limited, which may be related to the current contextual adaptability of AI voice interaction and the complexity of medical professional context.

Based on the above, this current study proposes the following practical suggestions: First, educators can build a "step-by-step" academic English training framework based on DeepSeek, such as focusing on terminology accumulation and sentence polishing at the primary stage, and introducing AI assistance in literature review writing and simulated academic debate at the advanced stage. Second, it is recommended that technology developers further optimize DeepSeek's medical vertical field functions, such as embedding interfaces of authoritative medical literature databases (such as PubMed), or developing special writing templates for medical case reports to enhance the subject-specificity of the tool. In addition, teachers need to balance the roles of AI tools and manual guidance in blended teaching to avoid students' over-reliance on technology and weakening critical thinking, such as guiding students to reflect on the revision process through the dual feedback mechanism of "AI initial evaluation + teacher review".

Future studies can be deepened from three directions: first, long-term tracking of the effect of AI-enabled teaching and observing the sustainable improvement of students' academic English ability; second, exploring the combination of DeepSeek and other technologies (such as virtual reality clinical scenarios) to strengthen the application context of medical English; third, comparative studies from a cross-cultural perspective, such as analyzing the differences in the benefits of medical students from different language backgrounds from AI tools. In short, AI technology represented by DeepSeek provides a new paradigm for medical academic English teaching, but its implementation needs to take into account educational laws, subject characteristics and technical feasibility in order to achieve true "empowerment" rather than "replacement".

VI. ACKNOWLEDGMENT

The current researchers sincerely thank all the medical students who participated in this experiment. It was their active cooperation and valuable feedback that provided real and reliable data support for the current study. The researchers would also like to thank the team members for their hard work in data collection and analysis, as well as the inspiration given by peer scholars in academic discussions. Most importantly, the current researchers are particularly grateful to the IJCRT editorial experts as well as those blind review experts for their constructive suggestions during the paper review process, which further improved the theoretical framework and method design of this study and promoted the successful completion of this research paper.

REFERENCES

[1] Cao, K. (2025). "DeepSeek Moment": A Political Analysis from Open Source Technology to Technological Justice [J/OL]. *United Front Studies.*

[2] Huang, X., Zou, D., Cheng, G., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society, 26*(1), 112-131.

[3]Kam Cheong Li, Billy Tak & Ming Wong. (2025). Agile-Blended Learning: An Innovative Technology-Mediated Approach to Distance Education. *SN Computer Science (2)*, 116-127.

[4] Kohnke, L., & Moorhouse, B. L. (2022). Adopting AI-based tools in English language teaching: A community of practice perspective. *RELC Journal, 53*(3), 370-384.

[5] Li, R., & Liu, M. (2021). AI-assisted medical English learning: A systematic review. *Journal of Medical Education and Curricular Development, 8*, 1-12.

[6] Liu, C., & Mo, S. (2025). Clarification of ideological security boundaries, risk monitoring and regulatory logic in generative artificial intelligence applications - Reflections based on DeepSeek, Manus, ChatGPT and Sora [J/OL]. *Journal of Chongqing University (Social Sciences).*

[7] Li, X., & Wang, Y. (2020). Challenges in academic English writing for Chinese medical students: A survey study. *English for Specific Purposes, 58*, 1-12.

[8] Li, Z., & Luo, X. (2025). Human resource reconstruction under the intelligent revolution: Research on the challenges, impacts and governance of DeepSeek and Manus-like generative artificial intelligence on the human resources market [J/OL]. *Journal of Chongqing University (Social Sciences).*

[9] Sarré, C., Grosbois, M., & Brudermann, C. (2021). Fostering accuracy in L2 writing: Impact of different types of automated feedback. *Computer Assisted Language Learning, 34*(5), 707-728.

[10] Wang, W., & Wang, D. (2025). The core characteristics, value transcendence and future direction of China's generative artificial intelligence DeepSeek [J/OL]. *United Front Studies.*

[11] Wang, Y., & Chen, N. S. (2023). AI-powered writing assistants in higher education: A mixed-methods study on student engagement and learning outcomes. *Journal of Educational Computing Research, 61*(1), 181-203.

[12] Xu, D., & Wang, B. (2025). A new paradigm of digital civilization: ecological construction of DeepSeek open source innovation [J/OL]. *United Front Studies.*

[13] Xu, W., & Ouyang, F. (2022). A systematic review of AI applications in language education: Opportunities and challenges. *International Journal of Educational Technology in Higher Education, 19*(1), 1-27.

[14] Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J. B., Yuan, J., & Li, Y. (2021). A review of AI-driven blended learning in post-COVID-19 era. *Educational Technology Research and Development, 69*(1), 217-228.