Enhancing Problem-Based Learning in Undergraduate Sports Medicine Education with DeepSeek AI: A Comparative Study

Yumei Wu, Jiahao Meng, Pan Liu, yilin Xiong, Shenghong Chen, Heng Yin, Shuhao Wang, Shuguang Gao

Submitted to: JMIR Medical Education
on: September 04, 2025

Table of Contents

Enhancing Problem-Based Learning in Undergraduate Sports Medicine Education with DeepSeek AI: A Comparative Study

Yumei Wu[1] BA; Jiahao Meng[1] MD; Pan Liu[1] BA; yilin Xiong[1] MD; Shenghong Chen[1] MD; Heng Yin[1] BA; Shuhao Wang[1] BA; Shuguang Gao[1] MD, PhD

[1] Department of Orthopaedics, Xiangya Hospital, Central South University, Changsha, Hunan, China. Changsha CN

Corresponding Author:
Shuguang Gao MD, PhD

Department of Orthopaedics, Xiangya Hospital, Central South University, Changsha, Hunan, China.
Department of Orthopaedics, Xiangya Hospital, Central South University, Changsha, Hunan, China
Changsha
CN

Abstract

Background: Problem-based learning (PBL) is a cornerstone of modern medical education, yet its effectiveness can be hindered by inconsistent student engagement and facilitator variability. The recent emergence of large language models (LLMs) like DeepSeek offers a potential to revolutionize pedagogical approaches by assisting in content creation and lesson design.

Objective: This study aimed to evaluate the impact of a DeepSeek-assisted PBL teaching model compared to traditional lecture-based teaching on undergraduate medical students' academic performance, learning efficiency, and satisfaction in a sports medicine course.

Methods: A quasi-experimental study was conducted with two randomly assigned classes of fourth-year medical students. The control group (n=33) received a traditional multimedia lecture on sports injuries. The experimental group (n=80) received a PBL session on the same topic, where the content was developed with the assistance of DeepSeek. Both groups completed a 20-item knowledge assessment and a 5-domain satisfaction survey immediately after the 45-minute session. The Mann-Whitney U test was used to compare outcomes due to non-normal data distribution.

Results: There was no significant difference in post-test scores between the traditional and DeepSeek-assisted PBL groups (82.27 vs 81.00, P=.64). Although the DeepSeek group completed the assessment faster (493.58 seconds vs 556.06 seconds), this difference was not statistically significant (P=.27). Overall satisfaction was high in both groups. Notably, the DeepSeek-assisted PBL group reported higher satisfaction rates in specific domains: 97.5% of students were satisfied with teaching resources (including AI-assisted tools) and instructor interaction.

Conclusions: The integration of DeepSeek AI into a PBL framework did not lead to superior academic performance in the short term compared to traditional methods. However, its potential to enhance learning efficiency and student satisfaction with interactive resources is promising. These findings suggest that AI-assisted PBL is a feasible and engaging pedagogical tool. Future research should explore long-term integration, optimized prompt engineering, and qualitative insights to fully leverage AI's potential in medical education.

(JMIR Preprints 04/09/2025:83291)
DOI: https://doi.org/10.2196/preprints.83291

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http
No. Please do not make my accepted manuscript PDF available to anyone. I understand that if I later pay to participate in <a href="https:/

Original Manuscript

Enhancing Problem-Based Learning in Undergraduate Sports Medicine Education with DeepSeek AI: A Comparative Study

Abstract

Background

Problem-based learning (PBL) is a cornerstone of modern medical education, yet its effectiveness can be hindered by inconsistent student engagement and facilitator variability. The recent emergence of large language models (LLMs) like DeepSeek offers a potential to revolutionize pedagogical approaches by assisting in content creation and lesson design.

Objective

This study aimed to evaluate the impact of a DeepSeek-assisted PBL teaching model compared to traditional lecture-based teaching on undergraduate medical students' academic performance, learning efficiency, and satisfaction in a sports medicine course.

Methods

A quasi-experimental study was conducted with two randomly assigned classes of fourth-year medical students. The control group (n=33) received a traditional multimedia lecture on sports injuries. The experimental group (n=80) received a PBL session on the same topic, where the content was developed with the assistance of DeepSeek. Both groups completed a 20-item knowledge assessment and a 5-domain satisfaction survey immediately after the 45-minute session. The Mann-Whitney U test was used to compare outcomes due to non-normal data distribution.

Results

There was no significant difference in post-test scores between the traditional and DeepSeek-assisted PBL groups (82.27 vs 81.00, *P*=.64). Although the DeepSeek group completed the assessment faster (493.58 seconds vs 556.06 seconds), this difference was not statistically significant (*P*=.27). Overall satisfaction was high in both groups. Notably, the DeepSeek-assisted PBL group reported higher satisfaction rates in specific domains: 97.5% of students were satisfied with teaching resources (including AI-assisted tools) and instructor interaction.

Conclusions

The integration of DeepSeek AI into a PBL framework did not lead to superior academic performance in the short term compared to traditional methods. However, its potential to enhance learning efficiency and student satisfaction with interactive resources is promising. These findings suggest that AI-assisted PBL is a feasible and engaging pedagogical tool. Future research should explore long-term integration, optimized prompt engineering, and qualitative insights to fully leverage AI's potential in medical education.

Keywords: artificial intelligence; problem-based learning; education, medical; educational technology; sports medicine; DeepSeek

Introduction

Problem-based learning (PBL) has been widely adopted as a core pedagogical strategy in medical education worldwide since its inception at McMaster University in the 1960s(1). This student-centered approach, which challenges learners to collaboratively solve complex, real-world clinical problems, aims to develop not only foundational knowledge but also critical thinking, self-directed learning, and teamwork skills essential for clinical practice(2, 3). Despite its theoretical benefits, the effective implementation of traditional PBL faces significant challenges. These include high dependency on skilled facilitators, considerable time and resource demands for case development, and variability in student participation and engagement, which can lead to inconsistent educational outcomes (4).

The landscape of undergraduate medical education, particularly in foundational courses such as

sports medicine, is often dominated by didactic lecture-based teaching. This model, while efficient for knowledge transmission to large cohorts, frequently struggles to foster deep conceptual understanding, clinical reasoning, and long-term knowledge retention among students(5). As medical knowledge continues to expand rapidly, there is a pressing need for innovative teaching methodologies that can enhance educational efficiency and effectiveness without disproportionately increasing the burden on instructors (6) .

The recent advent of advanced large language models (LLMs) presents a transformative opportunity to address these longstanding challenges. DeepSeek (DeepSeek Chat, DeepSeek Inc.), a powerful LLM developed in China, has demonstrated remarkable capabilities in natural language understanding, content generation, and contextual reasoning(7). Beyond its general applications, its potential in educational settings is increasingly being recognized. LLMs like ChatGPT have been explored for generating medical cases, simulating patient interactions, and providing personalized feedback(8, 9). Specifically, in a PBL context, AI can serve as a valuable assistant to instructors by rapidly generating nuanced case narratives, creating differential diagnoses, suggesting probing questions to stimulate discussion, and curating relevant learning resources, thereby enriching the entire PBL process(10).

However, despite the growing interest in AI-assisted education, there is a notable scarcity of empirical research that rigorously evaluates its impact in a structured classroom setting. Most existing studies are either exploratory, focusing on the feasibility of AI-generated content, or subjective, relying on student perceptions of utility(11, 12). Few have conducted controlled comparisons to assess the objective learning outcomes and efficiency gains of integrating an LLM like DeepSeek into a proven pedagogical model like PBL for medical students.

Therefore, this study aims to fill this gap by investigating the integration of DeepSeek into a PBL session for undergraduate sports medicine education. We conducted a quasi-experimental study to compare the effects of a DeepSeek-assisted PBL approach with those of traditional lecture-based teaching on students' academic performance, learning efficiency, and satisfaction. We hypothesize that while the AI-assisted approach may not significantly improve immediate test scores due to the short intervention time, it will enhance learning efficiency and student engagement, demonstrating its potential as a valuable tool for modernizing medical education.

Methods

Study Design and Participants

This quasi-experimental study compared a DeepSeek-assisted PBL approach with traditional lecture-based teaching in a single session on sports injuries. Two intact classes of fourth-year undergraduate medical students at Central South University were randomly assigned to either the intervention or control group. All participants had completed identical prerequisite coursework, ensuring baseline knowledge equivalence. Students who were absent or provided incomplete responses were excluded from the final analysis. The study was approved by the Institutional Review Board of Xiangya Hospital, and informed consent was obtained from all participants.

Intervention: DeepSeek-Assisted PBL Development

The same instructor delivered both sessions. For the control group (n=36), a traditional 45-minute multimedia lecture was prepared. Subsequently, the instructor used DeepSeek (DeepSeek Chat, DeepSeek Inc.) to reformat the lecture content into a PBL case scenario. The AI was prompted to generate a clinical case narrative, create guided discussion questions, and suggest diagnostic reasoning pathways based on the original learning objectives. The instructor then refined AI-generated materials into a structured PBL lesson plan.

Procedure

The control group received the traditional lecture. The experimental group (n=91) then participated in the revised, DeepSeek-assisted PBL session, which was also 45 minutes in duration. In the final five minutes of both sessions, all students completed an identical 20-question in-class assessment (covering basic knowledge and clinical application) and a 5-item satisfaction survey (Supplement).

Data Collection and Measures

Primary outcomes were academic performance (test score out of 100) and task completion time (seconds). Secondary outcomes were student satisfaction scores across five domains (clarity, interaction, resources, helpfulness, overall satisfaction), measured on a 5-point Likert scale (from 1= “Very dissatisfied” to 5= “Very satisfied”). Demographic data (age, gender) were also collected.

Statistical Analysis

Due to the non-normal distribution of data (assessed by Shapiro-Wilk test), non-parametric tests were employed. Mann-Whitney U tests were used to compare continuous variables (test scores, time) between groups, while Chi-square tests were used for categorical demographics. Effect sizes were calculated as rank-biserial correlation (r) for Mann-Whitney U tests. Statistical analyses were performed using R software (version 4.4.2), with a significance level set at $P < .05$.

Results

Participant Characteristics

After excluding absent students and incomplete responses, data from 33 students in the control group and 80 in the DeepSeek-assisted PBL group were analyzed. The two groups were comparable in terms of age (control: mean 21.42, SD=0.89; intervention: mean 21.71, SD=1.30; *P*=.42) and gender distribution (female: 63.64% vs. 63.75%; *P*=.99), indicating successful randomization.

Academic Performance and Efficiency

No significant difference was found in post-test scores between the traditional lecture group and the DeepSeek-assisted PBL group (82.27 ±14.53 vs. 81.00 ±15.64, respectively; U=1393, *P*=.64, r=0.05). Although students in the DeepSeek group completed the assessment faster on average (493.58 ±175.15 seconds vs. 556.06 ±236.06 seconds), this difference did not reach statistical significance (U=1146.5, *P*=.27, r=0.13). The distribution of scores and completion times is visualized in Figure 1.

Student Satisfaction

Overall satisfaction was high in both groups. However, the DeepSeek-assisted PBL group reported higher satisfaction in specific domains: 97.5% were satisfied with "teaching resources" and "instructor interaction," compared to 96.97% across all domains in the traditional group. A detailed breakdown of satisfaction ratings is presented in Table 1 and Figure 2.

Table 1: Participant Satisfaction Outcomes

Items		Very satisfied	Satisfied	Neutral	Dissatisfied	Very dissatisfied	Satisfaction (%)
1.How satisfied are you with the clarity and logical flow of this	Traditional	23	9	1	0	0	**96.97%**
	Deepseek-assisted PBL	54	20	6	0	0	**92.5%**

course?							
2.How satisfied are you with the instructor' s explanation and interaction (e.g., questions, discussions) ?	Traditional	23	9	1	0	0	**96.97%**
	Deepseek-assisted PBL	56	22	2	0	0	**97.5%**
3.How satisfied are you with the teaching resources used (e.g., PPT, case studies, AI tools)?	Traditional	26	6	1	0	0	**96.97%**
	Deepseek-assisted PBL	55	23	2	0	0	**97.5%**
4.To what extent did this course help your learning?	Traditional	23	9	1	0	0	**96.97%**
	Deepseek-assisted PBL	55	19	6	0	0	**92.5%**
5.How satisfied are you with the course overall?	Traditional	23	9	1	0	0	**96.97%**
	Deepseek-assisted PBL	56	19	5	0	0	**93.75%**

Discussion

Principal Findings

This quasi-experimental study evaluated the integration of the DeepSeek large language model into a problem-based learning session for undergraduate sports medicine education. The principal finding was that the DeepSeek-assisted PBL approach did not yield a statistically significant improvement in immediate knowledge acquisition compared to traditional lecture-based teaching, as measured by post-test scores. However, a promising and consistent trend was observed in learning efficiency, with the intervention group completing the assessment faster, albeit not at a conventional level of significance. Notably, students in the DeepSeek-assisted group reported exceptionally high satisfaction, particularly with the quality of teaching resources and the level of instructor interaction, suggesting that the AI-enhanced format positively influenced their perceptual learning experience.

Comparison with Prior Work

Our finding that AI assistance did not lead to superior test scores is consistent with some early research into generative AI in education, which suggests that these tools may not function as a direct catalyst for short-term knowledge gain but rather as a facilitative agent for the learning process

itself(13, 14). The trend toward improved efficiency (reduced completion time) aligns with the proposed value of LLMs in reducing cognitive load and streamlining information processing for learners. By providing well-structured cases and resources, the AI may have helped students navigate the problem-solving process more swiftly (15, 16).The high student satisfaction, especially regarding teaching resources and interaction, is a critical outcome. This strongly echoes the findings of Monzon & Hays (2025), who reported that AI-assisted learning environments significantly boost student motivation and engagement(10). The AI-curated content likely provided a novel, interactive, and visually coherent learning experience that resonated with students.

Limitations

Several limitations of this study must be acknowledged. First, the single-session, single-instructor design may limit the generalizability of the findings and introduces the potential for instructor bias, although this was mitigated by using the same instructor for both groups. Second, the relatively short duration of the intervention may have been insufficient to elicit measurable differences in knowledge retention, as the benefits of AI-enhanced PBL might be cumulative and more evident in long-term application. Third, the unequal group sizes, while reflective of real-world class allocations, may have reduced the statistical power to detect smaller effect sizes, particularly for the efficiency metric. Finally, the study measured immediate outcomes; the long-term impact on knowledge decay, clinical reasoning skills, and engagement remains unknown.

Future Directions

Future research should address these limitations and build upon our findings. Longitudinal studies with multiple sessions and integrated AI prompts are needed to investigate sustained effects on learning outcomes. Importantly, future work should incorporate qualitative methods, such as focus groups and think-aloud protocols, to gain a deeper understanding of how students interact with and perceive the AI-generated content, uncovering mechanisms behind the efficiency gains and high satisfaction(17). Furthermore, research should explore the concept of “prompt engineering” for educators-developing and validating optimal frameworks for instructors to use LLMs most effectively for pedagogical content creation(18). Finally, expanding this research to multi-institutional settings and different medical sub-disciplines would enhance the external validity of the findings.

Conclusions

In conclusion, this study demonstrates that the integration of the DeepSeek LLM into a PBL framework is a feasible and highly acceptable approach among undergraduate medical students. While it did not produce a significant improvement in short-term academic performance, its potential to enhance learning efficiency and student engagement is substantial. The high satisfaction ratings indicate that AI-assisted PBL can create a more engaging and resource-rich learning environment. These findings support the notion that AI should not be viewed as a replacement for traditional pedagogy or skilled facilitators, but rather as a powerful complementary tool to augment and refine the PBL process. As medical education continues to evolve, AI models like DeepSeek hold significant promise for scaling effective, engaging, and efficient educational experiences, empowering instructors and enriching student learning.

Abbreviations: LLM, Large Language Model; PBL, Problem-Based Learning

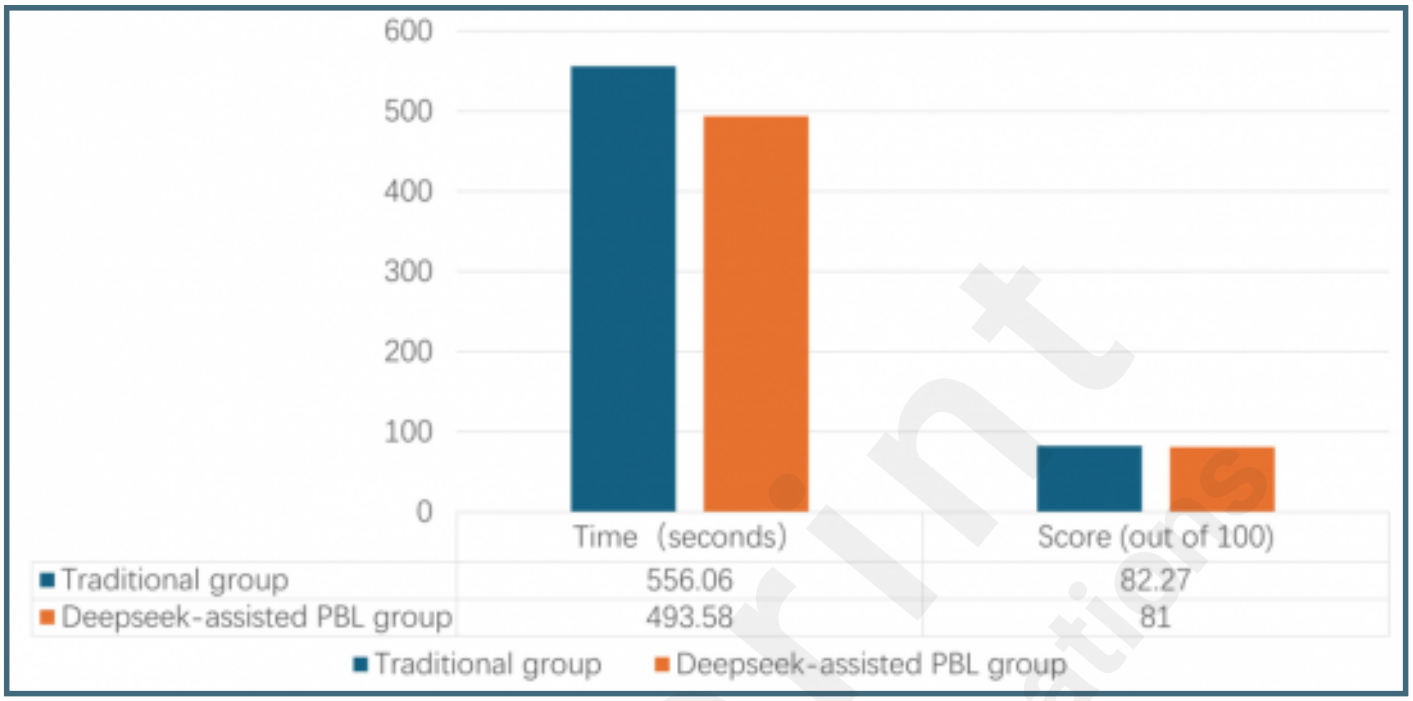
References

1. Mensour EA, Tran C, Li T, Mallawaarachchi I, Shaw JM, Blissett S. Evaluating the outcomes of problem-based learning in postgraduate medical education: a systematic review and meta-analysis. Can Med Educ J. 2025;16(1):89-99.
2. Atta IS, Alghamdi AH. The efficacy of self-directed learning versus problem-based learning for teaching and learning ophthalmology: a comparative study. Adv Med Educ Pract. 2018;9:623-630.
3. Forbes HM, Syed MS, Flanagan OL. The Role of Problem-Based Learning in Preparing Medical Students to Work As Community Service-Oriented Primary Care Physicians: A Systematic Literature Review. Cureus. 2023;15(9):e46074.
4. Dolmans D, Gijbels D. Research on problem-based learning: future challenges. Med Educ. 2013;47(2):214-218.
5. Taylor DC, Hamdy H. Adult learning theories: implications for learning and teaching in medical education: AMEE Guide No. 83. Med Teach. 2013;35(11):e1561-1572.
6. Wartman SA, Combs CD. Medical Education Must Move From the Information Age to the Age of Artificial Intelligence. Acad Med. 2018;93(8):1107-1109.
7. Normile D. Chinese firm's large language model makes a splash. Science. 2025;387(6731):238.
8. Divito CB, Katchikian BM, Gruenwald JE, Burgoon JM. The tools of the future are the challenges of today: The use of ChatGPT in problem-based learning medical education. Med Teach. 2024;46(3):320-322.
9. Jacobs C. Examining Multimodal AI Resources in Medical Education: The Role of Immersion, Motivation, and Fidelity in AI Narrative Learning. JMIR Med Educ. 2025;11:e72190.
10. Monzon N, Hays FA. Leveraging Generative Artificial Intelligence to Improve Motivation and Retrieval in Higher Education Learners. JMIR Med Educ. 2025;11:e59210.
11. Wang J, Fan W. The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: insights from a meta-analysis. Humanities and Social Sciences Communications. 2025;12(1):621.
12. Charles KA, Yousuf A, Chua HC, Matthews S, Harnett J, Hinton T. AI in action: Changes to student perceptions when using generative artificial intelligence for the creation of a multimedia project-based assessment. Eur J Pharmacol. 2025;998:177508.
13. Ahmed A, Kerr E, O'Malley A. Quality assurance and validity of AI-generated single best answer questions. BMC Med Educ. 2025;25(1):300.
14. Fan L, Deng K, Liu F. Educational impacts of generative artificial intelligence on learning and performance of engineering students in China. Sci Rep. 2025;15(1):26521.
15. Gkintoni E, Antonopoulou H, Sortwell A, Halkiopoulos C. Challenging Cognitive Load Theory: The Role of Educational Neuroscience and Artificial Intelligence in Redefining Learning Efficacy. Brain Sci. 2025;15(2).
16. Jiang T, Wu J, Leung SCH. The cognitive impacts of large language model interactions on problem solving and decision making using EEG analysis. Front Comput Neurosci. 2025;19:1556483.
17. Ravšelj D, Keržič D, Tomaževič N, Umek L, Brezovar N, N AI, et al. Higher education students' perceptions of ChatGPT: A global study of early reactions. PLoS One. 2025;20(2):e0315011.
18. Chen B, Zhang Z, Langrené N, Zhu S. Unleashing the potential of prompt engineering for large language models. Patterns (N Y). 2025;6(6):101260.

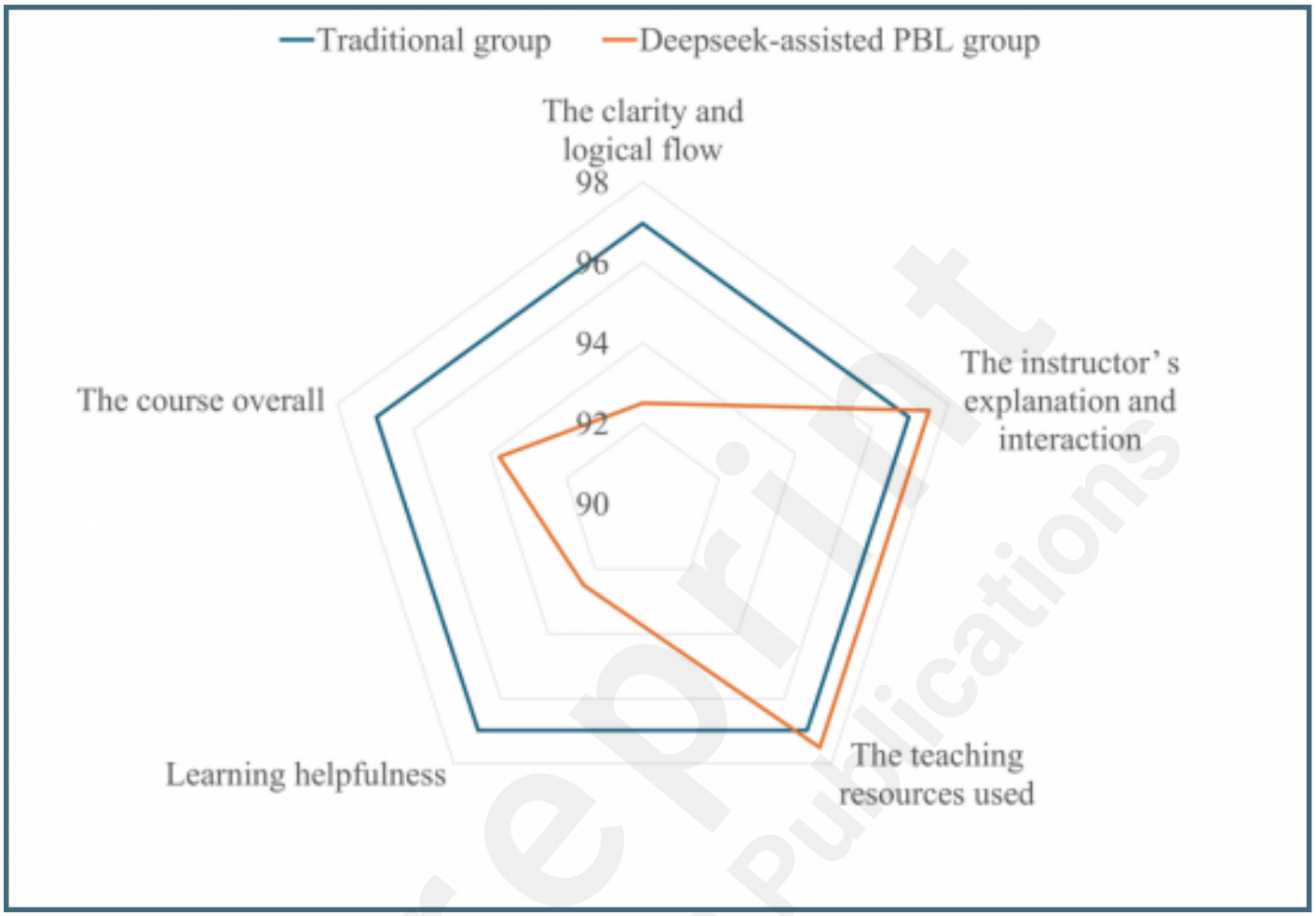
Supplementary Files

Figures

A radar chart comparing satisfaction rates (% satisfied/very satisfied) for each of the five survey items between the two groups.



A radar chart comparing satisfaction rates (% satisfied/very satisfied) for each of the five survey items between the two groups.



Related publication(s) - for reviewers eyes onlies

Survey questionnaire content.
URL: http://asset.jmir.pub/assets/62237664322f85496d0396403fed29c0.pdf