Integration of Artificial Intelligence in the Education of Future Teachers of Physics

UDC: 372.853 DOI: https://doi.org/10.15421/172590

Jalilova Sevinj
Ph.D., Assoc. Prof., https://orcid.org0000-0002-4753-2835, sevinjjalilova@yahoo.com
Azerbaijan State Pedagogical University (Baku, Azerbaijan)

Abstract.
The transformative potential of AI in education, is based on the architecture of transformers that are able to diagnose students' misconceptions with high accuracy and hence validity. Nevertheless, students still face persistent cognitive difficulties in learning physics, including misconceptions and gaps in background knowledge. Therefore, the hypothesis of the study was to hypothesize the feasibility of using ChatGPT as an additional pedagogical tool within a constructivist inquiry-based approach to improve conceptual understanding and address misconceptions in a thermodynamics course for freshman students who are future physics teachers.

The purpose of the study was to examine the effectiveness of using ChatGPT, a generative AI chatbot, as an additional pedagogical tool in a thermodynamics course for freshmen students, future physics teachers. The scientific novelty of the study lies in revealing the mechanism of improving students' conceptual understanding of thermodynamics principles, eliminating misconceptions in the knowledge component and creating an interactive learning environment through ChatGPT, opening new opportunities for improving educational practices in professional education. The results of the study showed that the experimental group of students significantly outperformed the control group in improving conceptual understanding and reducing qualitative misconceptions.

The results achieved emphasize the great potential of ChatGPT in STEM education within a constructivist learning environment to transform it. The practical significance of the research results is presented by a specific pedagogical toolkit that can act as an additional means of improving conceptual understanding and eliminating misconceptions in a thermodynamics course for freshman students - future physics teachers.

Keywords: STEM education, artificial intelligence, students, physics, ChatGPT

Інтеграція штучного інтелекту в освіту майбутніх вчителів фізики

Джалілова Севіндж
Азербайджанський державний педагогічний університет (Баку, Азербайджан)

Анотація.
Трансформаційний потенціал штучного інтелекту в освіті базується на архітектурі трансформаторів, здатних діагностувати помилкові уявлення студентів з високою точністю та, отже, валідністю. Тим не менш, студенти все ще стикаються з постійними когнітивними труднощами у вивченні фізики, включаючи помилкові уявлення та прогалини у фонових знаннях. Тому гіпотеза дослідження полягала в тому, щоб висунути гіпотезу про доцільність використання ChatGPT як додаткового педагогічного інструменту в рамках конструктивістського дослідницького підходу для покращення концептуального розуміння та усунення помилкових уявлень у курсі термодинаміки для студентів-першокурсників, які є майбутніми вчителями фізики.

Метою дослідження було вивчити ефективність використання ChatGPT, генеративного чат-бота на базі штучного інтелекту, як додаткового педагогічного інструменту в курсі термодинаміки для студентів-першокурсників, майбутніх вчителів фізики. Наукова новизна дослідження полягає у розкритті механізму покращення концептуального розуміння студентами принципів термодинаміки, усунення помилкових уявлень у складі знань та створення інтерактивного навчального середовища через ChatGPT, що відкриває нові можливості для вдосконалення освітньої практики у професійній освіті.

Результати дослідження показали, що експериментальна група студентів значно перевершила контрольну групу в покращенні концептуального розуміння та зменшенні якісних помилкових уявлень. Отримані результати підкреслюють великий потенціал ChatGPT у STEM-освіті в рамках конструктивістського навчального середовища для її трансформації. Практичне значення результатів дослідження представлено специфічним педагогічним інструментарієм, який може виступати додатковим засобом покращення концептуального розуміння та усунення помилкових уявлень у курсі термодинаміки для студентів-першокурсників – майбутніх вчителів фізики.

Ключові слова: STEM-освіта, штучний інтелект, студенти, фізика, ChatGPT

Стаття надійшла / Article arrived: 28.07.25 Схвалено до друку / Accepted: 26.08.25

Introduction.

Recent advances in artificial intelligence (hereinafter - AI) have been driven by significant developments in three critical areas of knowledge: evolution of neural network architectures, improvement of machine learning algorithms, and spread of deep learning methodologies (Hong, 2023, p. 45). The above-mentioned technological innovations have led to the emergence of a large variety of language models that have significantly expanded AI's ability to process, generate, and interact with natural language. Generative AI models, such as ChatGPT, Gemini and Claude, demonstrate impressive results in solving various tasks, including communication, content creation, complex problem solving and data analysis (Ali, Shamsan, Hezam & Mohammed, 2023, p. 42). It is natural to see the impact of the above developments on education, transforming pedagogical practices, supporting personalized learning approaches, and redefining the educational experience.

Building on these innovations, AI-based tools offer unprecedented opportunities to improve teaching and learning processes. They facilitate personalization of learning, provide instant feedback, and provide adaptive learning support tailored to the diverse needs of students (Xu & Ouyang, 2022, p. 59]. For example, generative AI models such as ChatGPT simulate human-like learning, clarify complex concepts, and engage students in interactive dialogues that foster deeper understanding of learning material (Lee, Yang & Shin, 2020, p. 339). Generative models of AI provide dynamic and adaptive support through real-time dialogic interaction, addressing misconceptions and knowledge gaps in STEM education.

Relevance. The transformative potential of AI in education, is based on the architecture of Transformers, which can diagnose students' misconceptions with high accuracy, and hence validity (Ahmed et al, 2022, p. 8). Nevertheless, students still face persistent cognitive difficulties in learning physics, including misconceptions and basic knowledge gaps. Addressing knowledge gaps is critical to stimulating academic progress and developing students' complex problem solving and critical thinking skills, which are increasingly required in STEM fields. The present study examines the effectiveness of using ChatGPT, a generative AI chatbot, as a supplemental pedagogical tool in a thermodynamics course for freshman students, future physics teachers.

The purpose of the study is to improve students' conceptual understanding of thermodynamics principles, eliminate misconceptions in the knowledge component, and create an interactive learning environment.

Analysis of previous studies and publications. While AI-based pedagogical tools are increasingly being introduced into the educational process, few studies have focused on their application in pedagogical specialties or their potential for solving subject-specific problems. This study aims to address the gaps by contributing to the discourse on the integration of AI in science education and offering a framework for the application of generative AI tools in educational settings.

Generative AI has significantly changed the education system by making learning personalized and adaptive by introducing context-aware responses and dynamic feedback to facilitate interactions with complex educational content and support various activities.

The effectiveness of generative AI tools largely depends on the quality of user prompts, as accurate input directly affects the relevance of the result. Despite their potential, generative AI systems often demonstrate limited contextual understanding, making deeper analysis difficult. However, their use in philosophical dialogues has demonstrated the ability to improve analysis-based discussions and facilitate meaningful participation in a variety of activities (Johnson, & Lee, 2020, p. 47). The results already achieved in research activities allow us to consider generative AI as a universal tool for solving persistent problems in STEM learning.

Thermodynamics, often considered an abstract and conceptually complex subject, has benefited from a variety of pedagogical strategies to improve learning outcomes and student engagement (Aktay, Gök & Uzunoğlu, 2023, p. 401). Constructivist approaches, such as the energy chain model, emphasize active learning and the integration of qualitative and quantitative reasoning to address persistent misconceptions in thermodynamics (Gottlieb, Kline & Schneider, 2023, p. 82).

Gamification strategies improve collaboration and engagement but may favor superficial interaction at the expense of deeper conceptual understanding (Park, Teo, T., Teo, A., Chang, Huang & Koo, 2023, p. 9).

Simulation-based learning environments promote learning and critical thinking by allowing students to visualize thermodynamic phenomena. The innovative tools offered allow working with variables to help learn concepts such as gas behavior, pressure-volume dependence, and particle interactions (Gayed, Carlon, & Oriola, 2022, p. 125). Computational tools additionally support active problem solving on complex topics including numerical simulations and equations of state, providing real-time feedback (Fryer, Coniam & Carpenter, 2020, p. 19). New technologies such as virtual reality (hereafter referred to as VR) and augmented reality (hereafter referred to as AR) have brought immersive elements to thermodynamics education. For example, Thermo VR includes embedded tests to simulate piston-cylinder systems, facilitating visualization and providing instant feedback (Luo & Yang, 2022, p. 7). These innovative teaching methods demonstrate the potential to make thermodynamics more accessible to students. Generative AI complements these strategies by bridging conceptual gaps, facilitating interactive learning, and empowering an educational process tailored to each student's individual abilities.

Materials and Research Methods. The present study utilized a quasi-experimental design with continuous

monitoring and evaluation of the impact of ChatGPT, a chatbot with generative AI, as a supplementary pedagogical tool in a thermodynamics course for undergraduate prospective physics teachers (n = 87). The participants were divided into two groups: a control group (CG, n = 44) that followed traditional teaching methods and an experimental group (EG, n = 43) that used ChatGPT in their teaching activities.

In addition, a continuous assessment approach was implemented using a survey on student acceptance of technology and improvement of their learning process, which allowed for the collection of data on student engagement, and the impact of ChatGPT on learning outcomes over the course of the semester. The comprehensive research design was intended to provide a detailed understanding of the potential of generative AI tools, specifically ChatGPT, to improve educational practices.

The main instrument for this study was the questionnaire "Thermodynamic Processes and the First and Second Principles of Thermodynamics" (STPFSL) in the short form STPFSL-Short. This multiple-choice questionnaire provides an objective assessment of students' understanding of the fundamental principles of thermodynamics and helps instructors evaluate the effectiveness of innovative methods of teaching thermodynamics. The STPFSL-Short consists of 33 items covering basic thermodynamic principles. An optional student survey was administered to obtain an assessment of student acceptance of AI as a pedagogical tool. The 16-item questionnaire combines a five-point Likert scale and multiple-choice questions that focus on the frequency and context of ChatGPT use, its role in addressing misconceptions, and its impact on learning processes in general. This approach provided the systematic data collection needed to assess how generative AI tools affect students' cognitive engagement, conceptual understanding, and acceptance of technology in STEM education.

The experiment was conducted during class sessions that were extended by 30 minutes each week, resulting in a total of 32 hours of class time per semester. Classes were held in a classroom with stable internet access and computers. Students used the free version of ChatGPT (versions 3.5 and later), accessed at through the OpenAI platform, as an additional pedagogical resource. An additional 30 minutes of each session was devoted to activities such as developing and refining prompts, analyzing AI-generated responses for accuracy and depth, and collaboratively engaging in instructor-led discussions of the results. These discussions addressed conceptual problems, clarified misconceptions, and explored the principles of thermodynamics in real-world scenarios. The instructor provided real-time feedback on the quality of problem-solving approaches, and collaborative discussions reinforced both theoretical understanding and practical application of thermodynamic principles.

Results.

After the experimental data were collected, statistical analysis was performed using R (version 4.4.1) and Microsoft Excel to process, visualize and interpret the data. The data were first checked for missing values and outliers, which were subsequently eliminated. Then, difficulty and discrimination scores for STPFSL were calculated according to recommendations based on classical test theory (Lee & Perret, 2022, p. 12788]. Prior to factor extraction, a parallel Horn analysis was performed to determine the optimal number of retained components to ensure that only components reflecting significant variance were selected. Promax rotation was then applied to account for potential correlations between factors. The Kaiser-Meyer-Olkin (KMO) criterion, Bartlett's sphericity criterion, and Cronbach's alpha for each aspect of the survey were found to be appropriate. The resulting subscale reliability coefficients ranged from 0.581 to 0.701. The overall Cronbach's alpha coefficient of 0.693 indicated moderate internal consistency. For the STPFSL, reliability was found to be acceptable, with values ranging from 0.618 to 0.690. In EG at the control stage of the experimental activity, reliability was higher (pre-test: 0.784, final test: 0.823) than in CG, reflecting improved internal consistency as a result of the pedagogical intervention.

Pre-test results showed no statistically significant differences between the control and experimental groups (Tabl. 1.). However, after the intervention, the EG showed higher final test scores (M = 22.53, SD = 3.03) compared to the CG (M = 18.60, SD = 1.90).The EG also showed higher learning scores (10.31 vs. 6.25 in the CG), indicating the potential of ChatGPT to improve conceptual learning in thermodynamics.

Tabl. 1.

Comparison of EG and CG learning outcomes

Stages	EG		CG	
	Mean value / Standard deviation	Range	Mean value / Standard deviation	Range
Stating	12,22 (3,15)	5,00-20,00	12,35 (2,64)	5,00-19,00
Control	22,53 (3,03)	15,00-28,00	18,60 (1,90)	13,00-22,00
Degree of effectiveness	10,31	5,00-14,00	6,25	3,00-9,00

The Mann-Whitney U-criterion confirmed a statistically significant difference in final test scores between CG and EG (U = 518.50, p < 0.001, Cohen's d = 0.614). In addition, Wilcoxon's sign rank criterion showed a significant improvement in within-group scores from the pre-test to the final test (W = 0, Z = -9.506, p < 0.001), with the EG showing significantly higher scores.

The chi-square criterion showed a significant decrease in the number of misconceptions in both groups, with the EG showing greater improvement overall. For example, misconceptions about isobaric ($\chi2(119) = 150.71$, p = 0.026) and adiabatic processes ($\chi2(119) = 149.18$, p = 0.032) decreased in the CG. In contrast, EG showed greater improvements, especially in irreversible processes ($\chi2(119) = 236.84$, p < 0.001) and entropy ($\chi2(119) = 186.49$, p < 0.001). These results suggest that ChatGPT is more effective in eliminating misconceptions in complex topics. Both groups showed marginal improvement on the topics of Reversible Processes and Motor Efficiency, indicating the need for more deliberate learning strategies. Additional measures such as targeted multi-step assignments and real-world applications may be needed to effectively learn these complex topics.

Tabl. 2. presents the results of the impact of ChatGPT on the learning process, indicating that there is a strong relationship between the frequency of ChatGPT use and its perceived effectiveness.

Tabl. 2.

Comparative analysis of students' perceptions of the effectiveness of ChatGPT as a function of frequency of use

Variables	Category	M	SD	U	P-value
Ease of use	HF	4,16	0,737	336	0,985
	HF	4,20	0,561		
Elimination of misconceptions	HF	3,93	0,495	240	0,037*
	HF	4,27	0,618		
Satisfaction with explanations	HF	3,87	0,352	284	0,755
	HF	4,16	0,562		
Overall impression	HF	4,33	0,617	338	0,287
	HF	4,22	0,517		
Compared to other methods	HF	4,20	0,596	336	0,357
	HF	4,13	0,516		

*Note: HF - high frequency; LF - low frequency, M - mean value, SD - standard deviation, U - Mann-Whitney U-criteria, * - statistical significance*

A series of Mann-Whitney U-criteria were conducted to compare ChatGPT users' perceptions on several parameters. A statistically significant difference was found on the usefulness parameter, where EG users rated ChatGPT as more useful (M = 4.27, SD = 0.594) compared to CG users (M = 3.93, SD = 0.495), U = 240.5, p = 0.037. For other parameters such as ease of use, elimination of misconceptions or overall impression, no significant differences were found. This suggests that the usefulness of ChatGPT remains stable regardless of the frequency of use.

Most students rated ChatGPT positively, with 31.67% rating it as "very positive" and 55% rating it as "positive", noting the user-friendly interface and interactive features that enhanced the learning process. In addition, 71.67% of students appreciated its informative help, emphasizing its value in supporting learning. However, responses to questions about overcoming misconceptions were more varied, indicating room for improvement, especially in dealing with complex misconceptions that are more difficult to address with superficial feedback.

In addition, 41.67% of students expressed a strong intention to recommend ChatGPT, and its anticipated future use emphasizes its potential as a valuable educational tool. Although the advantages of ChatGPT in terms of usability and content delivery are clear, the relationship between students' intention to recommend it and learning outcomes will provide a more complete picture of its overall effectiveness.

Conclusions.

The findings highlight the potential of AI-enabled generative tools to improve STEM education through personalized real-time feedback and active student engagement. The present study contributes to the understanding of the need for the use of ChatGPT in the professional preparation of educators. According to the results of the study, EG students achieved significantly higher learning outcomes (10.31 points compared to 6.25 points in the CG) and formed a more holistic understanding of the principles of thermodynamics. In addition, the study demonstrated that ChatGPT effectively addressed misconceptions in knowledge areas such as entropy, internal energy, and the second law of thermodynamics. EG students showed a greater reduction in misconceptions (e.g., regarding irreversible processes, p < 0.001) compared to CG students, as well as a marked improvement in knowledge of abstract topics such as PV diagrams and pressure-volume dependence. Also, based

on the results of the study, it can be confidently stated that ChatGPT has a positive effect on students' engagement in learning. For instance, 71.67% of the students appreciated ChatGPT for its ease of use and informative help.

The present study highlights the transformative potential of ChatGPT as an additional pedagogical tool for organizing and improving professional education. In conclusion, ChatGPT has significant potential to improve conceptual understanding and reduce misconceptions. However, its effectiveness depends on thoughtful implementation, continuous monitoring, and integration of additional pedagogical tools. The practical relevance of the research results stems from the expanded range of possibilities to utilize generative AI tools as a powerful complement to traditional teaching methods, making science and physics education accessible and engaging worldwide.

REFERENCES

Ahmed, A., Ampry, E., Komariah, A., Hassan, I., Thahir, I., Hussein, M., Ali, & Zafarani, P. (2022). Investigating the effect of using game-based learning on EFL learners' motivation and anxiety. *Education Research International,* P. 1-9.

Aktay, S., Gök, S., & Uzunoğlu, D. (2023). ChatGPT in education. *Türk Akademik Yayınlar Dergisi (TAY Journal)*, 7(2), 378-406.

Ali, J., Shamsan, M., Hezam, A., & Mohammed, T. (2023). Impact of ChatGPT on learning motivation: teachers and students' voices. *Journal of English Studies in Arabia Felix*, 2(1), 41-49.

Fryer, L. K., Coniam, D., & Carpenter, R. (2020). Bots for language learning now: Current and future directions. *Language Learning & Technology*, 24(2), 8-22.

Gayed, J. M., Carlon, M. K. J., & Oriola, A. M. (2022). Cross Exploring an AI-based writing assistant's impact on English language learners. Computers and Education: Artificial Intelligence, 3(1), 123-126.

Gottlieb, M., Kline, J. A., & Schneider, W C. (2023). ChatGPT and conversational artificial intelligence: Friend, foe, or future of research? *The American Journal of Emergency Medicine*, 70, 81-83.

Hong, W. (2023). The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research. *Journal Of Educational Technology and Innovation*, 3, 37-45.

Johnson, R., & Lee, M. (2020). The role of AI in modern physics teaching. *Journal of Educational Technology,* 12(3), 45-56.

Lee, I., & Perret, B. (2022). Preparing high school teachers to integrate AI methods in STEM classrooms, in Conference Paper. 36th AAAI Conference on Artificial Intelligence 2022. *Proceedings*, 22 Feb-1 Mar 2022, held virtually. P. 12783-12791

Lee, J. H., Yang, H. & Shin, D. (2020). Chatbots -technology for the language teacher. *ELT Journal, 2020*, 74(3), 338-344.

Luo, Q., & Yang, J. (2022). The Artificial Intelligence and Neural Network in Teaching. *Computational Intelligence and Neuroscience,* 1, 1-11.

Park, J., Teo, T., Teo, A., Chang, J., Huang, J. S. & Koo, S. (2023). Integrating Artificial Intelligence into Science Lessons: Teachers' Experiences and Views. *International Journal of STEM Education*, 10, 1-22.

Xu, W., & Ouyang, F. (2022). The application of AI technologies in STEM education: a systematic review from 2011 to 2021. *In t. J. STEM Educ.*, 9, 59