Revista Românească pentru Educație Multidimensională
ISSN: 2066-7329 | e-ISSN: 2067-9270
Abstracting & Indexing | Clarivate Analytics - Web of Science (WoS); EBSCO; Google Scholar; ICI Journals Master List - Index Copernicus; Ideas RePeC; Econpapers; Socionet; CEEOL; Philpapers; KVK; WorldCat; CrossRef; J-GATE

2025, Volume 17, Issue 3, pages: 313-328 | https://doi.org/10.18662/rrem/17.3/1024
Submitted: July 1, 2025 | *Accepted for publication:* August 4, 2025

The Influence of AI-assisted Tools on Engineering Project Outcomes

Alexandru DINU [1]

[1] Faculty of Electronics, Telecommunications and Information Technology, National University of Science and Technology "Politehnica" Bucharest, Bucharest, Romania;
alexandru.dinu2408@upb.ro

Abstract: *In recent years, large language models (LLMs) have gained considerable attention in academic environments, particularly for their potential to support student work. This paper investigates how the use of LLM tools influences graduate engineering students' performance and perception during project-based learning. The study was conducted over the course of one semester at a Romanian technical university and involved 60 students, divided into two groups: one using AI tools such as ChatGPT and one working with traditional resources only. A mixed-method approach was employed, including pre- and post-project questionnaires and a dual evaluation system involving both human and AI grading, based on a shared rubric. Results show a significant increase in perceived satisfaction and a reduction in the reported difficulty among students who had access to AI tools. Moreover, their average grades were higher and more consistent compared to those in the non-AI group. The study also highlights the alignment between human and AI-based assessment and the growing openness among students toward adopting generative tools in future academic work. These findings suggest that integrating LLMs into higher education may improve learning experiences, but also raise questions about critical thinking, fairness, and ethical use.*

Keywords: *AI in education; large language models; higher education; project-based learning; student perception; academic performance; ChatGPT.*

How to cite: Dinu, A. (2025). The influence of AI-assisted tools on engineering project outcomes. *Revista Românească pentru Educație Multidimensională, 17*(3), 313-328. https://doi.org/10.18662/rrem/17.3/1024

©2025 Published by In Manifest Network in partnership with LUMEN Publishing. This is an open access article under the CC BY-NC-ND license http://creativecommons.org/licenses/by-nc-nd/4.0/

Introduction

In recent years, large language models (LLMs) have rapidly entered higher education settings, offering students new ways to approach academic work. Their ability to assist with writing, programming, and summarizing complex content has been especially noted in both writing-focused and data-intensive disciplines (Nelson et al., 2025; Uppal & Hajian, 2025; Ioanid & Andrei, 2025). In Romanian engineering education, the presence of these tools is becoming increasingly relevant, yet little is known about their actual impact on student outcomes in structured academic contexts.

This paper aims to examine how the use of LLMs influences student performance and experience in project-based learning. It presents a comparative study involving 60 graduate engineering students from a Romanian technical university. These students were divided into two groups: one allowed to use generative AI tools such as ChatGPT 4o, and one restricted to traditional academic resources.

The main objectives are to:

1. compare students' perceptions of project complexity and satisfaction
2. measure performance differences using a standardized grading rubric
3. evaluate the consistency between human and AI grading.

The study hypothesizes that AI-assisted students will report higher satisfaction, lower perceived difficulty, and more consistent academic results. These findings are discussed in the context of ongoing debates about the integration of AI tools into higher education practices (Jacobsen & Weber, 2025; Montenegro-Rueda et al., 2023).

Literature review

The increasing use of LLMs such as ChatGPT has led to a significant body of research focused on their applications in educational contexts. These tools are often praised for their role in enhancing learning efficiency, aiding in academic writing, and supporting personalized feedback (Jošt et al., 2024; Sajja et al., 2024; Susnjak, 2023). Several systematic reviews have highlighted their positive impact across domains, including programming, statistics, and writing-intensive disciplines (Naznin et al., 2025; Ali et al., 2024).

However, the widespread use of generative AI in education also raises important ethical and pedagogical concerns. Some researchers argue

that reliance on these tools may reduce students' critical thinking and original contribution (Jauhiainen & Guerra, 2023; Borji, 2023; Cotton et al., 2023). Other studies highlight the potential for academic dishonesty, the difficulty in tracking AI-generated content, and the risk of shallow engagement with learning materials (Islam & Islam, 2023; Khlaif et al., 2025).

While the potential of LLMs in improving academic efficiency is often acknowledged, the impact on cognitive development and long-term learning is still debated (Gilson et al., 2023; Kim et al., 2025). These concerns are particularly relevant in engineering education, where solution strategies and analytical reasoning are key learning goals (Mirea et al., 2025; Mollick & Mollick, 2023).

Most studies to date are perception-based or theoretical in nature, focusing on attitudes and potential rather than measurable learning outcomes (Al-Abdullatif & Alsubaie, 2024; Jang, 2024; Rahman & Watanobe, 2023). Only a small number of empirical studies have directly compared student work produced with and without AI support in controlled settings (Fernández Cerero et al., 2025; OpenAI, 2023; Ertem-Eray & Cheng, 2025).

The present paper contributes to this underdeveloped area by conducting a real-time, semester-long comparative study. It builds on prior research in statistical analysis and independence testing (Dinu et al., 2020a; Dinu, 2023), and integrates cognitive models that frame AI tools as reasoning scaffolds (Walpole & Myers, 1989; Dinu & Vlad, 2014; Dinu et al., 2020b). The study also extends discussions on LLMs in Romanian and European education systems (Dinu et al., 2020; Hamal et al., 2022; Bucea-Manea-Ţoniş et al., 2022), positioning this research within ongoing efforts to align AI-based pedagogy with academic rigor and ethical standards (Bostan, 2025; Pisica et al., 2023; Gupta et al., 2025)

Methodology

This study employed a comparative, quasi-experimental design to assess the impact of LLMs on student experience and academic performance. The participants were 60 master's-level students enrolled in engineering programs at the National University of Science and Technology Politehnica Bucharest. They were randomly assigned to two equal groups: one group had full access to LLM tools such as ChatGPT 4o, while the other was restricted to using traditional academic resources, including textbooks, lecture slides, and conventional web search engines.

Prior to the project start, all students completed a pre-project questionnaire to evaluate their familiarity with AI tools, their expectations

regarding project difficulty, and their openness to AI-assisted learning. The semester project required students to solve real-world data analysis problems using statistical methods. Topics included financial forecasting, environmental data modeling, and public dataset exploration. Students were allowed to work individually or in pairs, but the resource access remained group-dependent throughout.

After submitting their projects, students completed a post-project questionnaire measuring perceived task difficulty, satisfaction with their work, and attitudes toward the tools used. A standardized grading rubric was developed, assigning 40% of the score to technical accuracy, 30% to problem-solving and critical thinking, and 30% to clarity of communication. All projects were independently evaluated by a human professor and by ChatGPT 4o, which received anonymized versions of the reports and applied the same grading criteria.

To ensure rigor, all data were anonymized prior to analysis. Statistical tests included paired t-tests, descriptive statistics, and chi-square tests. Grade distributions and inter-rater consistency between human and AI evaluations were compared. The study complied with institutional ethical standards, and participants provided informed consent before participating in the research process.

Results

The comparative analysis of perceived project difficulty between the AI-assisted and traditional groups revealed several consistent and statistically significant trends, illustrated in Figure 1. Prior to initiating their semester projects, students in the NON-AI group generally anticipated a moderate level of difficulty. Most of them expected the task to be "Neutral," with a smaller proportion selecting "Easy" or "Difficult." Only a few anticipated it to be either "Very Easy" or "Very Difficult."

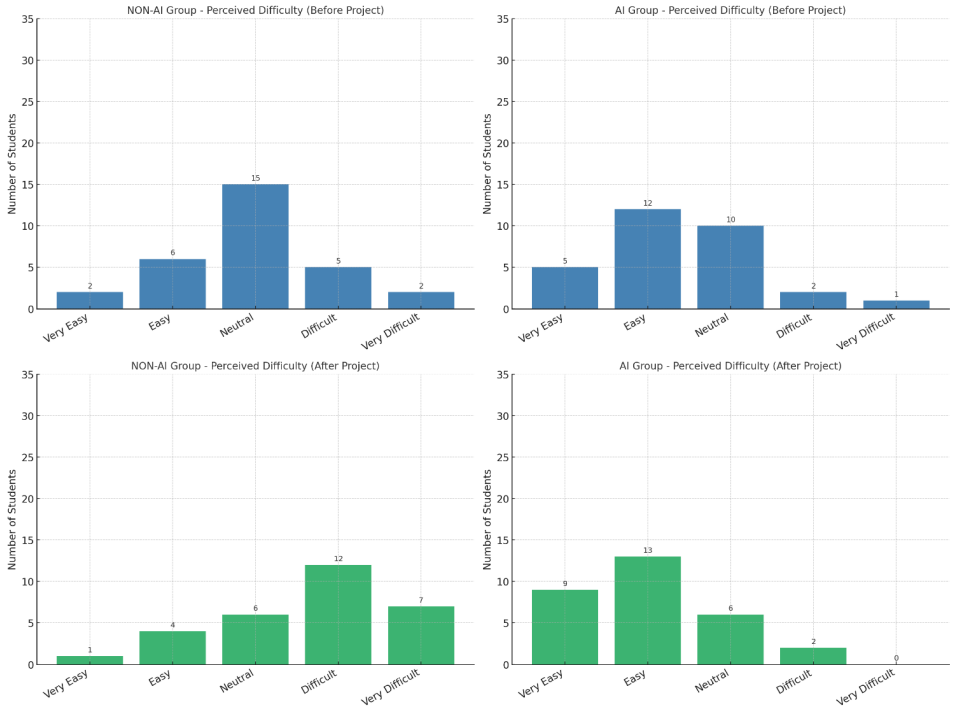
In contrast, students in the AI-assisted group demonstrated a more optimistic perspective at the outset. As shown in the same figure, most of them considered the project to be "Easy" or "Neutral," with very few expecting substantial difficulty.

By the end of the semester, these perceptions shifted in opposite directions. In the NON-AI group, the reported difficulty increased notably. As shown in the lower left quadrant of Figure 1, many students now described the project as "Difficult" or "Very Difficult," indicating that the workload exceeded their initial expectations. On the other hand, students in the AI group showed the reverse tendency: a substantial portion of them rated the project as "Very Easy" or "Easy," with almost no students

reporting high levels of difficulty after completing the task. This pattern suggests that the availability of AI tools reduced the cognitive and procedural load of the academic work, contributing to a more favorable overall experience.

These results support the initial hypothesis that incorporating LLM tools such as ChatGPT 4o into academic tasks lowers perceived project complexity and facilitates a more accessible learning experience.

Figure 1. Comparison of perceived project difficulty before and after project completion for NON-AI and AI-assisted groups. The top row shows students' expectations before starting the project, while the bottom row shows their perceptions after completion.

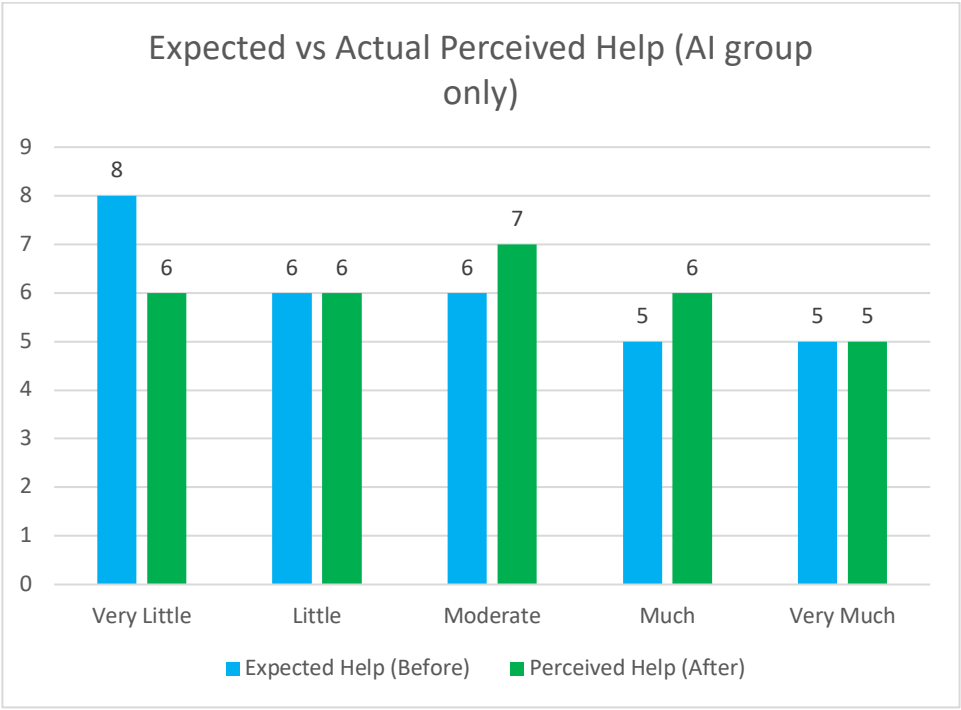


Source: author's own concept

In addition to these findings, the study analyzed the gap between expectations and actual experiences regarding AI support, focusing exclusively on the AI-assisted group. As shown in Figure 2, students initially expected the AI tools to help them "Much" or "Very Much." Very few respondents selected "Moderately" or "Little." After project completion, the perceived level of help increased even further: more students selected "Very Much," while responses in the lower categories dropped. This confirms that

the students not only validated their initial expectations but, in many cases, found the AI support to exceed what they had anticipated.

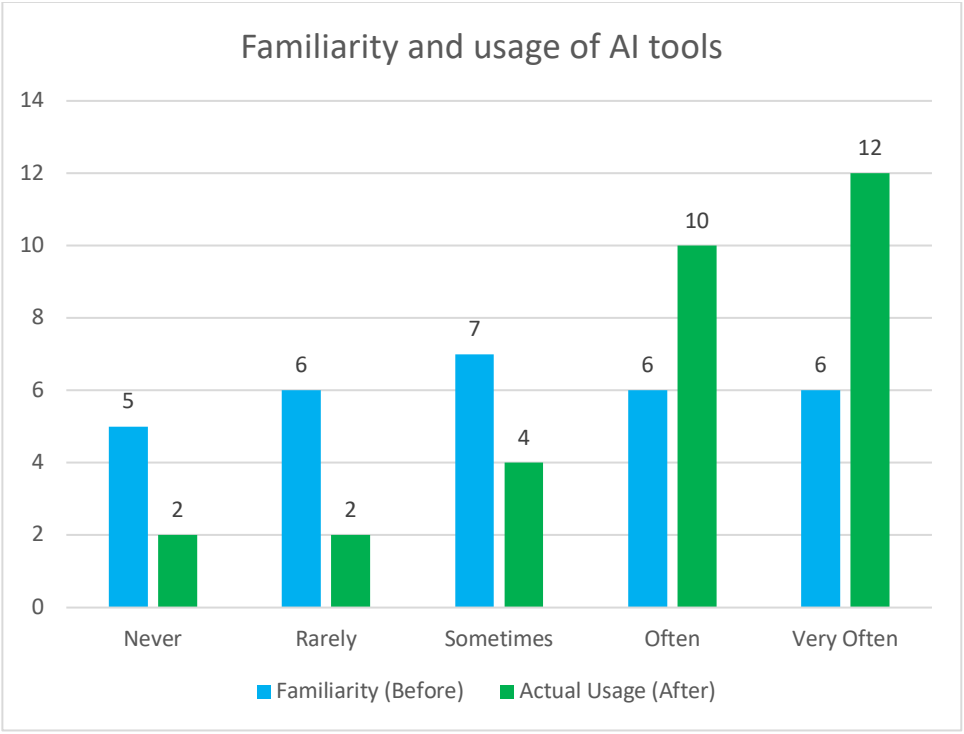
Figure 2. Comparison between expected and perceived help from AI tools in the AI-assisted group.



Source: author's own concept

Another dimension explored was the evolution of familiarity and usage of AI tools during the project. Figure 3 provides a comparison between reported familiarity before the project and actual frequency of use during the semester. At the beginning, many students reported limited interaction with such tools, indicating they had used them "Rarely" or "Sometimes," with only a few reporting frequent usage. However, after the project, most students declared they had used AI tools "Often" or "Very Often," while the number of those reporting rare use dropped substantially. This rapid increase in usage, despite initial unfamiliarity, suggests that students adapted quickly once they were encouraged and permitted to integrate AI into their workflow.

Figure 3. Comparison between students' familiarity with AI tools before the project and their actual usage during the project within the AI-assisted group.



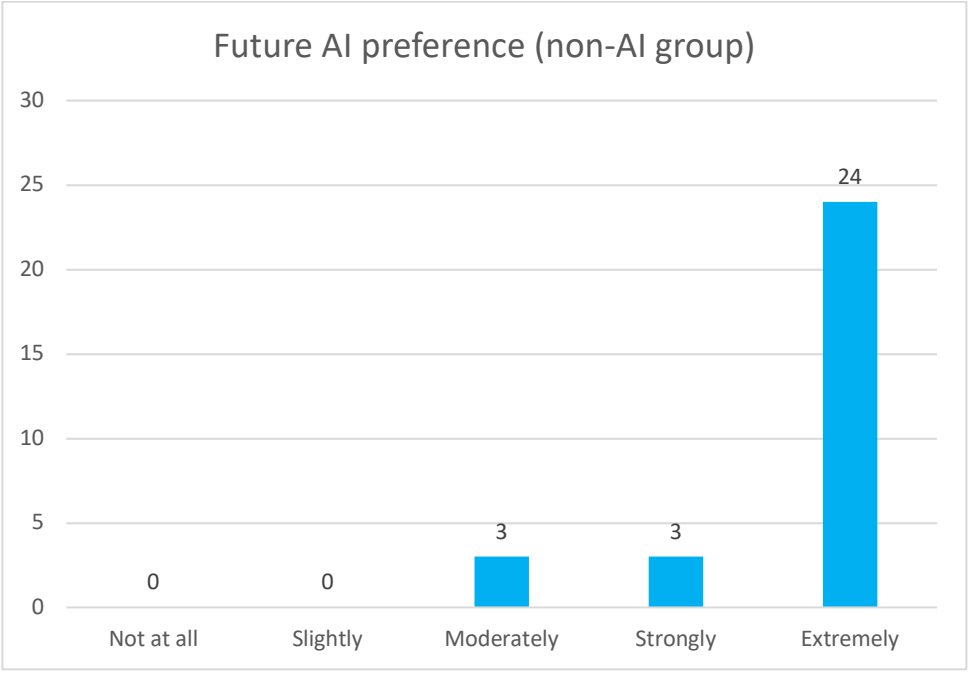
Source: author's own concept

The attitudes of the NON-AI group regarding future use of AI tools were also surveyed. As shown in Figure 4, most of these students indicated a clear interest in adopting such tools in subsequent academic projects. While a few expressed neutral or cautious opinions, the dominant trend reflected a positive inclination toward using AI resources in the future. This suggests that even those who were initially restricted from using LLMs came to recognize their practical advantages when faced with complex, open-ended academic tasks.

To further validate the impact of AI-assisted work, a comparative analysis of the final grades was conducted. The evaluation was performed independently by a human professor and by ChatGPT 4o, using an identical scoring rubric. As presented in Table 1, students in the AI-assisted group obtained an average grade of 8.93 from the professor and 9.10 from the AI evaluator. In contrast, students in the NON-AI group received an average of 7.67 and 8.07, respectively. These differences were statistically significant

according to two-sample t-tests, with p-values under 0.000001 for human grading and under 0.001 for AI grading.

Figure 4. Preference to use AI tools in future projects among students who worked without AI in the current semester.



Source: author's own concept

Table 1. Comparison of student performance (grades out of 10) for two groups: those using AI tools (1-30) and those without AI assistance (31-60), and two graders (human and AI).

Student ID	AI/NON-AI	Grade Prof	Grade GPT	Student ID	AI/NON-AI	Grade Prof	Grade GPT
1	AI	9	10	31	NON-AI	8	8
2	AI	10	9	32	NON-AI	7	8
3	AI	10	10	33	NON-AI	6	7
4	AI	9	9	34	NON-AI	9	9
5	AI	8	10	35	NON-AI	9	10
6	AI	9	10	36	NON-AI	9	10
7	AI	8	9	37	NON-AI	7	5
8	AI	10	9	38	NON-AI	7	7
9	AI	9	10	39	NON-AI	6	7
10	AI	10	10	40	NON-AI	9	9
11	AI	8	9	41	NON-AI	7	8
12	AI	10	9	42	NON-AI	7	7
13	AI	10	9	43	NON-AI	7	6
14	AI	9	9	44	NON-AI	8	7
15	AI	8	8	45	NON-AI	7	8
16	AI	8	8	46	NON-AI	9	9
17	AI	9	9	47	NON-AI	7	10
18	AI	9	9	48	NON-AI	7	9
19	AI	9	9	49	NON-AI	9	9
20	AI	8	8	50	NON-AI	8	9
21	AI	9	10	51	NON-AI	8	9
22	AI	9	10	52	NON-AI	6	7
23	AI	9	10	53	NON-AI	9	9
24	AI	9	9	54	NON-AI	8	6
25	AI	10	9	55	NON-AI	9	9
26	AI	10	10	56	NON-AI	9	9
27	AI	9	9	57	NON-AI	7	7
28	AI	9	9	58	NON-AI	7	7
29	AI	9	9	59	NON-AI	7	7
30	AI	8	8	60	NON-AI	7	7

Source: author's own concept

In addition to the mean differences, the standard deviation of grades was lower in the AI-assisted group (0.55 for professor, 0.44 for ChatGPT), suggesting that their output was not only of higher quality but also more consistent. Internal comparisons between the two evaluators (professor and AI) within each group did not reveal significant discrepancies: $p = 0.36$ for the AI-assisted group and $p = 0.18$ for the NON-AI group. These findings indicate a high level of agreement between human and AI-based grading when a common evaluation framework is applied.

Overall, the results confirm that students using LLM tools experienced enhanced academic performance, reduced perceived difficulty, and a greater sense of support, without compromising assessment objectivity or consistency.

Limits and discussions

While the present study provides empirical support for the integration of large language models (LLMs) in higher education, its conclusions must be interpreted in light of several limitations. First, the research was conducted in a single academic institution in Romania, which may limit the generalizability of the results to other geographical or institutional contexts. Although the university has a strong technical profile, differences in academic culture, digital literacy, and institutional openness to AI tools may affect how similar interventions would function elsewhere.

From an epistemological perspective, the study is based on a pragmatic understanding of learning outcomes, focusing on students' performance, perceived difficulty, and satisfaction. It does not investigate in detail the cognitive processes behind these outcomes, nor does it explore long-term learning retention. Further research using qualitative interviews or longitudinal designs would be needed to capture deeper insights into how AI tools influence critical thinking, conceptual understanding, or intellectual autonomy.

Methodologically, the research relies on self-reported perceptions and standardized rubrics for evaluation. Although the use of both human and AI grading adds robustness, subjectivity in self-assessment remains a factor. Moreover, the AI tool (ChatGPT 4o) was used as both an intervention and an evaluator, which, despite careful anonymization and consistency in rubric design, raises questions about circularity and tool bias.

Additionally, the study design did not involve full randomization of participants into treatment and control groups. While efforts were made to balance the cohorts in terms of academic background and program affiliation, the quasi-experimental nature of the group assignment may

introduce potential selection bias. Future implementations should consider randomized controlled trials to strengthen causal inference and rule out pre-existing group differences.

Furthermore, while the study focused on ChatGPT 4o as the primary AI tool, no comparative analysis was performed involving alternative LLMs or generative platforms (Gemini, Claude, Deep Seek, etc.). Considering the growing ecosystem of AI tools—each with distinct affordances, limitations, and user interfaces—a more comprehensive approach would involve multiple LLMs to better understand variation in educational effectiveness.

Another limitation lies in the lack of qualitative inquiry into students' cognitive engagement during the project. Although questionnaires provided valuable self-reported insights, complementary data from interviews, think-aloud protocols, or focus groups could offer a deeper understanding of how students interact with AI tools, how they process feedback, and whether such tools scaffold or bypass critical reasoning.

Finally, although the literature review is comprehensive, it leans heavily on sources from Western academic contexts. Expanding the citation base to include research from the other geographies would offer a more balanced perspective on how generative AI is perceived, regulated, and adopted in diverse cultural and infrastructural settings.

Despite these limitations, the findings can contribute meaningfully to the international literature on educational AI. The use of a real-life, semester-long intervention, with direct comparison between AI-assisted and non-assisted student groups, represents a relatively rare empirical approach. Most previous studies have focused on perceptions, opinions, or hypothetical use of AI (Al-Abdullatif & Alsubaie, 2024; Fernández Cerero et al., 2025). In contrast, this study operationalized LLM support in a controlled setting and evaluated its outcomes through both human and algorithmic lenses.

The results suggest that the inclusion of AI tools may enhance academic performance and learning experiences when thoughtfully integrated into project-based curricula. Moreover, the demonstrated alignment between human and AI grading contributes to ongoing discussions about the potential role of LLMs in academic assessment (Mollick & Mollick, 2023; Kim et al., 2025). These findings are relevant not only for Romanian institutions, but also for educators and researchers worldwide who seek to reconcile traditional pedagogical goals with emerging technological affordances.

The originality of this paper lies in its empirical, comparative design and its dual-evaluator approach. By measuring both perception and performance, and by involving AI both as a tool and as an evaluator, the study provides a layered analysis of how LLMs function in real educational scenarios. This design can serve as a model for future research exploring the educational utility and ethical implications of AI in higher education systems across different disciplines and cultural contexts.

Conclusions

The main objective of this study was to examine the influence of large language models on the academic experience and performance of graduate engineering students in a project-based learning environment. Through a semester-long comparative design involving two student groups—one using AI tools such as ChatGPT 4o, and one working with traditional resources—the study aimed to assess perceived difficulty, academic satisfaction, and learning outcomes.

The findings confirm the initial hypothesis: students who had access to AI tools reported lower perceived difficulty, higher satisfaction, and achieved statistically higher and more consistent grades. These results were supported by both human and AI-based grading, which showed a high degree of alignment. Furthermore, the research indicates that students quickly adapted to using LLMs and expressed a strong preference to continue using them in future academic work. Even students who were initially in the control group, and thus excluded from AI use during the project, expressed openness to adopting such tools in the future.

Based on these outcomes, the study concludes that the integration of generative AI into structured academic tasks can enhance student experience and performance without compromising evaluation standards. At the same time, the findings point to the need for further investigation into the long-term cognitive and ethical implications of widespread LLM use in education.

These conclusions reinforce the relevance of developing pedagogical frameworks that responsibly incorporate AI technologies, particularly in technical disciplines where problem-solving and data analysis are central. The research thus contributes both empirical evidence and methodological insight to the broader discourse on educational innovation in the era of artificial intelligence.

Acknowledgments

This research was supported by the Executive Unit for Financing Higher Education, Research, Development and Innovation (UEFISCDI) and the National University of Science and Technology "Politehnica" Bucharest, through project PN-IV-P7-7.1-PTE-2024-0505.

The author expresses his gratitude to the students who participated in this study and to the faculty colleagues who contributed feedback during the design and analysis phases.

All procedures involving human participants were conducted in accordance with institutional ethical standards. Informed consent was obtained from all participants, and no personal data were collected or processed.

The author declares no conflict of interest related to this study.

References

Al-Abdullatif, A. M., & Alsubaie, M. A. (2024). ChatGPT in learning: Assessing students' use intentions through the lens of perceived value and the influence of AI literacy. *Behavioral Sciences, 14*, 845. https://doi.org/10.3390/bs14090845

Ali, D., Fatemi, Y., Boskabadi, E., Nikfar, M., Ugwuoke, J., & Ali, H. (2024). ChatGPT in teaching and learning: A systematic review. *Education Sciences, 14*, 643. https://doi.org/10.3390/educsci14060643

Borji, A. (2023). A categorical archive of chatgpt failures. *arXiv preprint arXiv:2302.03494*. https://arxiv.org/abs/2302.03494

Bostan, I. (2025). Consolidating the role of AI in the economy and society: Combating the deepfake phenomenon through strategic and normative approaches—The case of Romania in the EU context. *Economies, 13*, 129. https://doi.org/10.3390/economies13050129

Bucea-Manea-Ţoniş, R., Kuleto, V., Gudei, S. C. D., Lianu, C., Ilić, M. P., & Păun, D. (2022). Artificial intelligence potential in higher education institutions: Enhanced learning environment in Romania and Serbia. *Sustainability, 14*, 5842. https://doi.org/10.3390/su14105842

Cotton, D., Cotton, P., & Shipway, J. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 60*(1), 15–26.https://doi.org/10.35542/osf.io/mrz8h

Dinu, A. (2023). Decoding the relationship between language, entropy, and cognitive functioning. *Journal of Educational Sciences and Psychology, 63*(2), 137–151. https://doi.org/10.51865/JESP.2023.2.11

Dinu, A., & Vlad, A. (2014). The compound tent map and the connection between Gray codes and the initial condition recovery. *Scientific Bulletin – UPB, 76*(1), 17–25.https://www.scientificbulletin.upb.ro/rev_docs_arhiva/reze8c_641132.pdf

Dinu, A., Vlad, A., Hanu, B., & Mitrea, A. (2020a). The statistical independence for words in printed Romanian language. In *Proceedings of the International Conference on Communications (COMM)* (pp. 319–324). https://doi.org/10.1109/COMM48946.2020.9142045

Dinu, A., Vlad, A., Hanu, B., & Mitrea, A. (2020b). Beginning and end of sentence word digrams for printed Romanian language. In *Proceedings of the 15th International Conference "Linguistic Resources and Tools for Natural Language Processing"* (pp. 54–61). Bucharest. https://profs.info.uaic.ro/~consilr/2020/proceedings2020.html

Ertem-Eray, T., & Cheng, Y. (2025). A review of artificial intelligence research in peer-reviewed communication journals. *Applied Sciences, 15*, 1058. https://doi.org/10.3390/app15031058

Fernández Cerero, J., Montenegro-Rueda, M., Román Graván, P., & Fernández Batanero, J. M. (2025). ChatGPT as a digital tool in the transformation of digital teaching competence: A systematic review. *Technologies, 13*, 205. https://doi.org/10.3390/technologies13050205

Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., &Chartash, D. (2023). How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. *JMIR Medical Education, 9*, e45312. https://doi.org/10.2196/45312

Gupta, O., Marrone, S., Gargiulo, F., Jaiswal, R., &Marassi, L. (2025). Understanding social biases in large language models. *AI, 6*, 106. https://doi.org/10.3390/ai6050106

Hamal, O., El Faddouli, N.–E., Harouni, M. H. A., & Lu, J. (2022). Artificial intelligence in education. *Sustainability, 14*, 2862. https://doi.org/10.3390/su14052862

Ioanid, A., & Andrei, N. (2025). ChatGPT in education: Challenges in local knowledge representation of Romanian history and geography. *Education Sciences, 15*, 511. https://doi.org/10.3390/educsci15040511

Islam, A. N., & Islam, M. N. (2023). ChatGPT in education: Opportunities and challenges. *Authorea.* https://doi.org/10.22541/au.167712329.97543109/v1

Jacobsen, L. J., & Weber, K. E. (2025). The promises and pitfalls of large language models as feedback providers: A study of prompt engineering and the

quality of AI-driven feedback. *AI, 6*, 35. https://doi.org/10.3390/ai6020035

Jang, M. (2024). AI literacy and intention to use text-based GenAI for learning: The case of business students in Korea. *Informatics, 11*, 54. https://doi.org/10.3390/informatics11030054

Jauhiainen, J. S., & Guerra, A. G. (2023). Generative AI and ChatGPT in school children's education: Evidence from a school lesson. *Sustainability, 15*, 14025. https://doi.org/10.3390/su151814025

Jošt, G., Taneski, V., & Karakatič, S. (2024). The impact of large language models on programming education and student learning outcomes. *Applied Sciences, 14*, 4115. https://doi.org/10.3390/app14104115

Khlaif, Z. N., Alkouk, W. A., Salama, N., & Abu Eideh, B. (2025). Redesigning assessments for AI-enhanced learning: A framework for educators in the generative AI era. *Education Sciences, 15*, 174. https://doi.org/10.3390/educsci15020174

Kim, H.–S., Kang, J.–W., & Choi, S.–Y. (2025). ChatGPT vs. human journalists: Analyzing news summaries through BERTScore and moderation standards. *Electronics, 14*, 2115. https://doi.org/10.3390/electronics14112115

Mirea, C.–M., Bologa, R., Toma, A., Clim, A., Plăcință, D.–D., &Bobocea, A. (2025). Transforming learning with generative AI: From student perceptions to the design of an educational solution. *Applied Sciences, 15*, 5785. https://doi.org/10.3390/app15105785

Mollick, E. R., & Mollick, L. (2023). Using AI to implement effective teaching strategies in classrooms: Five strategies, including prompts. *Wharton School Research Paper*. https://doi.org/10.2139/ssrn.4391243

Montenegro-Rueda, M., Fernández-Cerero, J., Fernández-Batanero, J. M., & López-Meneses, E. (2023). Impact of the implementation of ChatGPT in education: A systematic review. *Computers, 12*, 153. https://doi.org/10.3390/computers12080153

Naznin, K., Al Mahmud, A., Nguyen, M. T., & Chua, C. (2025). ChatGPT integration in higher education for personalized learning, academic writing, and coding tasks: A systematic review. *Computers, 14*, 53. https://doi.org/10.3390/computers14020053

Nelson, A. S., Santamaría, P. V., Javens, J. S., & Ricaurte, M. (2025). Students' perceptions of generative artificial intelligence (GenAI) use in academic writing in English as a foreign language. *Education Sciences, 15*(5), 611. https://doi.org/10.3390/educsci15050611

OpenAI. (2023). *GPT-4 technical report*. https://openai.com/research/gpt-4

Pisica, A. I., Edu, T., Zaharia, R. M., & Zaharia, R. (2023). Implementing artificial intelligence in higher education: Pros and cons from the perspectives of academics. *Societies, 13*, 118. https://doi.org/10.3390/soc13050118

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences, 13*, 5783. https://doi.org/10.3390/app13095783

Sajja, R., Sermet, Y., Cikmaz, M., Cwiertny, D., & Demir, I. (2024). Artificial intelligence-enabled intelligent assistant for personalized and adaptive learning in higher education. *Information, 15*, 596. https://doi.org/10.3390/info15100596

Susnjak, T. (2023). ChatGPT: The end of online exam integrity? *Education Sciences, 13*(2), 89. https://doi.org/10.3390/educsci14060656

Uppal, K., & Hajian, S. (2025). Students' perceptions of ChatGPT in higher education: A study of academic enhancement, procrastination, and ethical concerns. *European Journal of Educational Research, 14*(1), 199–211. https://doi.org/10.12973/eu-jer.14.1.199

Walpole, R. E., & Myers, R. H. (1989). *Probability and statistics for engineers and scientists* (4th ed.). Macmillan.