The impact of ChatGPT on student performance in higher education

Nimród Mike[1], Krisztina Karsai[2], Gábor Orbán[1], Alexandra Bubelényi[1], Csaba Norbert Nagy[1] and Gábor Polyák[3]

[1] *Corvinus University of Budapest, Fővám tér 8, 1093, Budapest, Hungary*
[2] *University of Szeged, Dugonics tér 13, 6720, Szeged, Hungary*
[3] *Eötvös Lóránd University, Múzeum krt. 6-8, 1088 Budapest, Hungary*

Abstract
This study investigates ChatGPT's impact on effectiveness, efficiency, and problem-solving among higher education students in law, business informatics, and media and communication. Involving 304 students divided into experimental (using ChatGPT) and control groups for an open-book test, the research aimed to assess efficiency benefits. Contrary to expectations, ChatGPT did not improve performance across disciplines. However, business informatics students completed tests faster, suggesting a nuanced effect on efficiency.

Keywords
ChatGPT, Higher Education, Large Language Model (LLM), Generative AI (GAI).

1. Introduction

Artificial Intelligence (AI) has significantly influenced education, with tools like ChatGPT enhancing student efficiency, problem-solving, and understanding in higher education [1], [4]. While AI's educational potential is widely recognized, its specific effects on student performance remain underexplored [9], [14]. Existing studies highlight AI's role in personalized learning and effective teaching strategies [4], [34], but there is a lack of empirical research comparing ChatGPT to traditional search tools like Google across various disciplines [10]. This study aims to fill this gap by examining ChatGPT's impact on student efficiency and problem-solving in higher education. Previous research, such as studies in Ghana, generally focused on AI's educational benefits [14].

We expand on this by exploring ChatGPT's nuanced effects across different academic disciplines, which has been underrepresented in the literature [22], [3], [21]. Our goal is to

Proceedings EGOV-CeDEM-ePart conference, September 1-5, 2024, Ghent University and KU Leuven, Ghent/Leuven, Belgium
EMAIL: nimrod.mike@uni-corvinus.hu (A.1); karsai.krisztina@szte.hu (A.2); orban.gabor@stud.uni-corvinus.hu (A.3); alexandra.bubelenyi@stud.uni-corvinus.hu (A.4); csaba.nagy2@stud.uni-corvinus.hu (A.5); polyak.gabor@btk.elte.hu (A.6)
ORCID: 0009-0005-7485-7271 (A.1); 0000-0003-4625-9774 (A.2); 0000-0001-9747-4373 (A.6)

provide insights that are relevant for both academic and policy-making contexts, helping institutions leverage AI in education frameworks. We hypothesize that ChatGPT will improve task efficiency but not necessarily enhance response accuracy compared to traditional methods [16], [28]. This is based on the idea that AI tools streamline information access and research processes, enhancing efficiency [26].

Our study aims to inform educational policies on integrating AI tools like ChatGPT to enhance teaching and learning [2]. Using a mixed-methods approach, we reviewed literature and conducted an experiment with 304 Hungarian students, comparing ChatGPT use in an open-book test to traditional methods. This research not only evaluates ChatGPT's impact on academic performance but also explores broader implications for digital governance and policy, reflecting growing interest in AI's societal role [24]. Our findings add empirical data on ChatGPT's efficiency and problem-solving benefits, advocating for a tailored approach to AI integration in education [13], [35]. By detailing ChatGPT's effects across disciplines, we provide valuable insights for educators and policymakers interested in the future of AI in education [6], [18].

2. Methodology

We employ a mixed-methods approach, combining qualitative and quantitative data, to assess ChatGPT's impact on student performance. A thorough literature review informed our research plan, emphasizing the value of diverse data collection techniques in capturing the multifaceted effects of AI tools in education [15], [25].

2.1. Literature review process

To identify seminal publications within AI-related educational literature, we used mathematical and statistical tools to identify key studies. We focused on articles with a high number of authors and citations to ensure a comprehensive and diverse perspective. This approach was chosen to capture a wide range of viewpoints and a strong academic impact. We selected papers using the following formula:

$$x > Q_3 + 1.5 * IQR \quad (1)$$

where x is the observed value, Q_3 is the third quartile, and IQR is the interquartile range. This statistical method helped us identify upward outliers, suggesting complexity and broad acceptance in the scholarly community. We then created a directed graph to visualize the interconnectedness of over sixteen hundred publications. This analysis revealed seven subgraphs, with a dominant interconnected subgraph indicating a shared knowledge base. Key publications within this subgraph were identified through their incoming edge degree and PageRank, highlighting the most influential knowledge hubs.

Our analysis provided a comprehensive overview of current research on ChatGPT in higher education. We noted a significant study conducted in Ghana involving a flipped classroom setup with 125 students, which demonstrated ChatGPT's positive impact on critical, creative, and reflective thinking [14]. Recent research extensively explores the

integration and impact of generative AI (GAI) tools like ChatGPT across various educational and professional contexts.

Recent works emphasize the importance of AI literacy and practical learning, suggesting research directions to prepare students for a society increasingly powered by GAI [9]. The technological evolution of digital writing and summarization frameworks has also been a focus, illustrating how ChatGPT can facilitate more advanced educational applications [22], [21].

Additionally, the transformative potential of ChatGPT and its implications for redefining academic "originality" are explored in studies that discuss how ChatGPT challenges traditional notions of academic integrity, urging educational institutions to adapt their policies accordingly [3], [23]. Other papers discuss necessary adaptations to teaching and assessment practices considering GAI, highlighting how educational frameworks need to evolve to accommodate these technologies [5], [35].

In addition to educational impacts, several studies address the broader ethical and societal implications of AI tools like ChatGPT. Some works examine the benefits and ethical concerns associated with using ChatGPT for scientific writing, noting the need for ethical guidelines in its application [1]. There is also analysis of varying public sentiments towards GAI, indicating a diverse range of opinions on its integration into society [24]. Foundational work and ethical considerations provide critical context for understanding the complex implications of deploying AI in educational and professional settings [6], [13].

Overall, the literature review emphasized the need for nuanced research into ChatGPT's impact across different educational contexts, aligning with our study's objectives to explore these effects in Hungarian universities. Recent research underscores the extensive application of ChatGPT in enhancing educational practices and highlights the importance of addressing the ethical and societal impacts of integrating AI tools in education.

2.2. Exploratory interviews

To contribute to our literature review, we conducted semi-structured interviews with four expert instructors from law, business informatics, and media studies. These interviews provided qualitative insights into the experiences and perceptions of educators regarding AI tools like ChatGPT.

Primary evidence from these interviews includes direct quotes and specific observations about the impact and challenges of integrating AI tools in higher education. For example, one instructor noted, "ChatGPT has significantly enhanced the speed at which students can gather initial research, but it lacks depth in more specialized areas," reflecting a common sentiment among participants. Supporting information includes additional context provided by the instructors, such as their perspectives on how AI tools are transforming teaching methodologies and academic integrity.

The interview protocol involved a set of guiding questions designed to explore various aspects of AI integration, including how do they perceive the role of ChatGPT in enhancing student learning and research capabilities; what are the main benefits and challenges they have encountered in using ChatGPT in their teaching practices, if they do so; and how do they address concerns regarding academic integrity and plagiarism in the context of AI tools.

Instructors expressed enthusiasm about ChatGPT's potential for idea generation but highlighted concerns about its limitations in specialized research areas. Discussions centered around the challenges of detecting AI-generated text and the implications for academic integrity, suggesting that university policies, rather than new legislation, should address these issues. This viewpoint emphasizes the need for institutional guidelines tailored to AI tools' unique challenges.

The interviews also revealed a consensus that academic tasks should evolve to incorporate AI tools, shifting the focus from memorization to critical thinking and problem-solving. Instructors proposed revising assignments to explicitly include the use of ChatGPT, thus teaching students how to effectively utilize AI tools. They emphasized that while AI might impact various professions, jobs requiring personal interaction would remain largely unaffected [24]. These insights suggest a direction for developing new educational strategies that integrate AI tools in a way that is both effective and ethically sound. As a conclusion, the interviews gave useful insights into how AI tools like ChatGPT are being used in education. They demonstrated the need to adjust our educational practices to include these tools in a way that keeps academic integrity intact and continues to develop critical thinking skills among students.

2.3. Pre-study preparations

We recruited undergraduate students from various Hungarian universities, including those studying law, business informatics, and media studies. Initially, 415 students enrolled in the study, providing a diverse sample that represents key academic disciplines relevant to our research. Participants were recruited through university email lists, ensuring a wide outreach.

To establish a baseline for the research, participants completed an assessment designed to evaluate their initial knowledge and attitudes towards ChatGPT. The assessment included tasks to measure critical thinking and problem-solving strategies, such as analyzing case studies, solving logical puzzles, and summarizing complex texts. Participants had 30 minutes to complete the assessment, with the average completion time recorded at 21.5 minutes.

Based on their performance in the pre-test, participants were divided into an experimental group (which utilized ChatGPT) and a control group (which relied on traditional methods). The assignment process involved balancing scores using a weighted formula that favored complex tasks, calculated as follows:

$$y^{student} = \sum_{task=1}^{n} score_{task}^{student} * \frac{1}{std_{task}} \quad (2)$$

2.4. Test designs

Developed in close collaboration with domain experts, the tests for law studies, business informatics, and media and communication Studies were structured to critically evaluate the efficacy of AI-assisted learning tools like ChatGPT within higher education. Each discipline's test consisted of 30 questions, employing a uniform format that included true/false, multiple-choice, fill-in-the-blank, and matching questions. This approach was

chosen to test a broad spectrum of knowledge and application skills across different educational domains, with the curriculum serving as the basis for question selection in each specific area.

2.5. Proprietary data collection tool

To gather comprehensive data, we developed ExamEye, a specialized browser extension that captured student interactions with ChatGPT and traditional search engines during controlled tests. ExamEye prioritized ethical standards and privacy, activating only within the testing platform and automatically ceasing recording upon test completion.

ExamEye provided a rich dataset, tracking participants' digital activity throughout the test environment. It recorded browsing activity, source type, and engagement with ChatGPT, including prompt crafting and response evaluation. This allowed us to distinguish between the use of ChatGPT and traditional research methods and analyze internet usage patterns. The tool's design ensured that only relevant data was collected, minimizing any potential privacy concerns.

Participants provided informed consent and were fully briefed on the use of ExamEye. Privacy safeguards included anonymizing data and restricting monitoring to the test environment only. The tool deactivated immediately upon test completion, ensuring that data collection adhered to ethical standards and protected participant privacy. Participants were informed about the data being collected and how it would be used, ensuring transparency and compliance with ethical guidelines.

2.6. Data collection and analysis

Data was collected using ExamEye, which monitored student interactions with AI tools during the tests. The data included usage patterns, time spent on tasks, and the nature of the interactions with ChatGPT. Statistical analysis involved comparing test completion times and accuracy between the experimental and control groups using t-tests and variance analysis to assess differences.

We employed independent sample t-tests to compare the means of test completion times and accuracy between the two groups, assessing whether the differences were statistically significant. Where variances were unequal, Welch's t-test was used to ensure robust results. Additionally, we performed sensitivity analyses to account for any variations in baseline performance that could influence the outcomes. This comprehensive approach ensured that our analysis was rigorous and reliable, providing clear insights into the impact of ChatGPT on student performance.

3. Results

The research hypothesis posits that the average score of students in the experimental group will be the same as that of the control group, with significantly better test times per stratum, suggesting that the use of ChatGPT will primarily affect efficiency. To confirm or refute these hypotheses, we first need to examine the significance of the differences between the

expected values of the test score (percentage) and test time (test_interval_s) variables for each stratum of the experimental and control groups separately.

The first step in hypothesis testing is to generate appropriate null and alternative hypotheses based on our assumptions. We pair our null hypothesis that the experimental group's outcome is identical to that of the control group on a stratified basis with partially overlapping alternative hypotheses, because the one-tailed alternative tests have a higher power of test than the two-tailed alternative test at the same level of significance, so that by comparing the p-value of each test we can draw a more accurate conclusion.

In order to test our hypothesis, we performed two independent sample t-tests to analyze the disparity of means. Prior to conducting these hypothesis tests, we ensured that all fundamental assumptions of the t-test were satisfied. This involved assessing the expected normality within small sample sizes (although the t-test remains robust to non-normality with sufficiently large sample sizes due to the central limit theorem) and confirming equal variances between the control and experimental groups within each stratum. In cases where significant differences in variances were observed, we opted for Welch's t-test over Student's t-test, as it accommodates both unequal variances and sample sizes. Once we confirmed that all test criteria were met, we proceeded with conducting the t-tests for each individual stratum.

Table 1. Showing hypothesis testing results for the difference in percentages between an experimental group (%[e]) and a control group (%[c]) across three test types: Communication, Informatics, and Law. It lists the null hypothesis (H0) that there is no difference, alternative hypotheses (H1) for non-equality or inferiority, and their corresponding p-values.

Test Type	H0	H1	p-value
Communication	μ(%[e]) - μ(%[c]) = 0	μ(%[e]) - μ(%[c]) != 0	0.052320
Informatics	μ(%[e]) - μ(%[c]) = 0	μ(%[e]) - μ(%[c]) != 0	0.609823
Law	μ(%[e]) - μ(%[c]) = 0	μ(%[e]) - μ(%[c]) != 0	0.968851
Communication	μ(%[e]) -μ(%[c]) = 0	μ(%[e]) - μ(%[c]) < 0	0.026160
Informatics	μ(%[e]) - μ(%[c]) = 0	μ(%[e]) - μ(%[c]) < 0	0.304912
Law	μ(%[e]) - μ(%[c]) = 0	μ(%[e]) - μ(%[c]) < 0	0.484425

Based on the findings, neither the informatics test nor the law test exhibited significant differences between the control and experimental groups in terms of both magnitude and direction. Thus, the observed variances could be attributed to random sampling, indicating that the discrepancies observed cannot be generalized to the broader participants. However, concerning the communication test, while the two-sided test indicated the difference as insignificant, the one-sided test, which is more sensitive to directional differences, deemed the gap significant with a p-value of 2.612%.

Table 2. Statistical Analysis of Mean Test Time Differences in Law, Communication, and Informatics.

Test Type	H0	H1	p-value
Communication	μ(t[e]) - μ(t[c]) = 0	μ(t[e]) - μ(t[c]) < 0	0.047682
Informatics	μ(t[e]) - μ(t[c]) = 0	μ(t[e]) - μ(t[c]) < 0	0.000179
Law	μ(t[e]) - μ(t[c]) = 0	μ(t[e]) - μ(t[c]) < 0	0.260312
Communication	μ(t[e]) - μ(t[c]) = 0	μ(t[e]) - μ(t[c]) != 0	0.095363
Informatics	μ(t[e]) - μ(t[c]) = 0	μ(t[e]) - μ(t[c]) != 0	0.000357
Law	μ(t[e]) - μ(t[c]) = 0	μ(t[e]) - μ(t[c]) != 0	0.520625

In the table provided, the test results regarding the mean test time difference are outlined. For the law test, despite observing a decrease in test time, the t-test deemed the difference as statistically insignificant in both magnitude and direction. As for the communication test, its p-values hovered around the boundary of significance, with the two-sided test accepting the null hypothesis while the one-sided test rejected it. This suggests a potential influence of ChatGPT on efficiency that warrants further investigation with a larger sample size. Conversely, the informatics test exhibited a clear and significant reduction in test time within the experimental group, indicating that ChatGPT significantly enhanced test completion efficiency in this discipline.

The results of the t-tests are attributed to the extent of ChatGPT usage and browser usage. It should be highlighted that 26 students from the experimental group and 23 from the control group could not be observed, in addition 70 students from the experimental group chose to use ChatGPT. Within this, a total of 1518 prompts were observed, with the middle 50% of students using between 10 and 32 prompts for the test. The prompts were often a specific copying of the questions, with rewriting the prompt for an incorrect answer being more common than rewriting the answer. On average, students spent 32.41 seconds writing the prompt, reading the answer and regenerate it if needed.

The distribution of ChatGPT usage across different test types reveals distinct patterns. For the communication test, 72% of students did not use ChatGPT, while 28% did. In the informatics test, 35% of students did not use ChatGPT, compared to 65% who did. For the law test, the majority of students, 81%, did not use ChatGPT, with only 19% utilizing it. This data highlights that ChatGPT usage was highest among informatics students and lowest among law students. Students in business informatics used ChatGPT extensively, which likely contributed to significantly better test times in this discipline. Focus time, indicating how long a student was actively engaged with the browser, was also measured. For business informatics, the average ChatGPT focus time was 6.45 minutes per test session, compared to 2.39 minutes for law and 2.30 minutes for communication. This suggests that the efficiency gains observed in business informatics may be attributed to the extensive use of ChatGPT and the structured nature of the tasks in this field.

The students who took the law and media and communication test in the experimental group also preferred to browse, using Google. In the case of the experimental group of the media and communication test, there was also little browsing activity, presumably explaining the lower scores in the experimental group.

Figure **1**. Average query count by participants

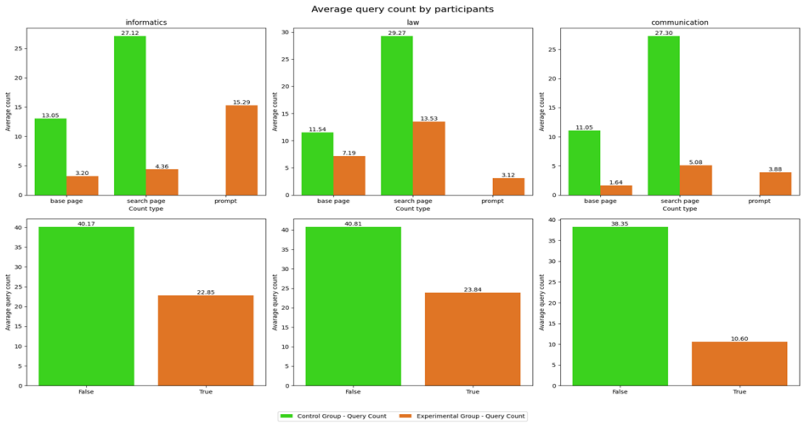

Figure 1 illustrates the reduction in internet usage among the experimental group, particularly notable in the field of business informatics. Despite lower ChatGPT usage in law and media and communication, there was still a noticeable decline in browser activity. This trend may stem from potential overconfidence induced by ChatGPT and decreased reliance on internet searches.

Upon scrutinizing the significance of deviations within the control group, we found it imperative to discern the extent to which differences could be accounted for by group composition and what remained as intervention-induced effects. Employing the two-fold Oaxaca-Blinder decomposition, we segregated the expected differences in test durations into explained and unexplained effects. Following the decomposition, it became evident that in the business informatics test, where the deviation was notably significant, intervention led to an average reduction in test completion time by nearly 6 minutes among students in the experimental group. Conversely, in the communication test, where the deviation approached significance, the composition effect exerted a stronger influence, leaving only 95 seconds unexplained by differences in group compositions. This highlights the differential impact of ChatGPT across disciplines and suggests that the tool may be particularly beneficial in structured fields like business informatics, where efficiency gains are more readily observable.

4. Discussion

4.1. Efficacy and effectiveness

The discussion is concerning both the literature review and the results. Our analysis indicates that while ChatGPT's integration into higher education research does not significantly enhance the accuracy of student responses, it notably improves the efficiency of completing tasks. This finding, endorsed by other researchers, suggests ChatGPT's

potential to boost student productivity by facilitating quicker research and problem-solving [16], [28]. The varied impact of ChatGPT across disciplines underscores the importance of a tailored approach to AI integration in education, cautioning against its unchecked use [19]. The varied impact of ChatGPT across disciplines highlights the need for a context-specific approach to its integration. While it can streamline certain academic tasks, it's essential to recognize its limitations in fostering critical thinking and creativity, which are vital for comprehensive learning. This insight calls for a balanced integration strategy, positioning AI tools as adjuncts to, rather than replacements for, traditional learning methods [17], [10]. Ensuring that AI tools enhance rather than diminish the development of key academic competencies is crucial for their effective use in educational settings.

4.2. Personalized learning

One of the standout benefits of integrating AI tools such as ChatGPT in educational settings is the facilitation of personalized learning [26]. AI's adaptive learning technologies can tailor educational content to meet the individual needs of students, considering their learning pace, preferred learning styles, and areas of difficulty. This personalized approach not only enhances student engagement but also addresses unique challenges, improving overall learning outcomes. Furthermore, the use of ChatGPT and similar technologies in education supports a more interactive and responsive learning environment. Unlike traditional educational resources, these AI tools can provide instant feedback and clarification, fostering a more dynamic and engaging learning experience [2]. The ability to provide real-time adjustments to the learner's needs significantly contributes to the effectiveness of the learning process, making learning more accessible and efficient.

4.3. Overreliance on technology

The dual potential of AI in education presents both opportunities for enhancement and risks of dependency. In their paper, researchers highlight that reliance on AI for learning can lead to a reduced development of critical thinking skills, even when AI explains its reasoning [32]. This underscores the need for a balanced approach in utilizing AI tools to maintain the integrity of educational processes. Educators must foster an environment where AI is used to complement, not replace, critical engagement and independent problem-solving skills.

4.4. Lack of deep-learning and critical analysis.

Reflecting on the challenges of implementing 'deep learning' in education, 'deep learning' here means thorough understanding and mastery of information, not just superficial details or problem-solving skills. It involves grasping core concepts, integrating new knowledge, and applying it broadly. Literature highlights skepticism about AI's role in fostering such deep learning, raising concerns over reliance on AI-generated answers and superficial content engagement [29]. Therefore, while AI tools like ChatGPT can offer quick information, they should be used to support and enhance deep learning by helping students gain a lasting understanding and effectively apply knowledge [22]. Educators should leverage AI to add context and reinforce concepts, ensuring it aids in active, meaningful learning.

5. Limitations

Our study has several constraints. Firstly, the participant selection from various Hungarian universities may limit the generalizability of our findings beyond this specific context [7]. Secondly, variations in students' prior experiences with ChatGPT or Google, study habits, and resource access could introduce biases that affect our result [18]. Additionally, the use of the ExamEye browser extension may have caused the Hawthorne effect, potentially altering students' genuine interactions with AI tools due to their awareness of being observed [8]. Lastly, the short study duration may not capture ChatGPT's long-term effects on educational outcomes, highlighting the need for future longitudinal research [30].

6. Conclusion

This study assessed ChatGPT's impact on higher education students in law, business informatics, and media and communication. While test scores did not improve, ChatGPT did speed up test completion times, with varying effects across disciplines.

References

[1] Altmäe, S., Sola-Leyva, A., & Salumets, A. (2023). Artificial intelligence in scientific writing: a friend or a foe? Reproductive BioMedicine Online, 47(1), 3-9. https://doi.org/10.1016/j.rbmo.2023.04.009

[2] Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. SSRN Electronic Journal.

[3] Bansal, G., Chamola, V., Hussain, A. et al. Transforming Conversations with AI—A Comprehensive Study of ChatGPT. Cogn Comput (2024). https://doi.org/10.1007/s12559-023-10236-2

[4] Bhattacharya, P., & Nakhare, S. (2019, November 1). Exploring AI-enabled Intelligent Tutoring System in the Vocational Studies Sector in UAE. https://doi.org/10.1109/itt48889.2019.9075093

[5] Bower, M., Torrington, J., Lai, J.W.M. et al. How should we change teaching and assessment in response to increasingly powerful generative Artificial Intelligence? Outcomes of the ChatGPT teacher survey. Educ Inf Technol (2024). https://doi.org/10.1007/s10639-023-12405-0

[6] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners. https://arxiv.org/abs/2005.14165

[7] Carminati, L. (2018). Generalizability in Qualitative Research: A Tale of Two Traditions. Qualitative Health Research, 28, 2094-2101.

[8] Chen, L., Vander Weg, M. V., Hofmann, D., & Reisinger, H. (2015). The Hawthorne Effect in Infection Prevention and Epidemiology. Infection Control & Hospital Epidemiology, 36(12), 1444–1450.

[9] Chiu, T. K. F. (2024). Future research recommendations for transforming higher education with generative AI. Computers and Education: Artificial Intelligence, 6, 100197. https://doi.org/10.1016/j.caeai.2023.100197
[10] Cotton, D., Cotton, P A., & Shipway, J R. (2023, March 13). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International, 1-12. https://doi.org/10.1080/14703297.2023.2190148
[11] Curry, N., Baker, P., & Brookes, G. (2024). Generative AI for corpus approaches to discourse studies: A critical evaluation of ChatGPT. Applied Corpus Linguistics, 4(1), 100082. https://doi.org/10.1016/j.acorp.2023.100082
[12] Einarsson, H., Lund, S. H., & Jónsdóttir, A. H. (2024). Application of ChatGPT for automated problem reframing across academic domains. Computers and Education: Artificial Intelligence, 6, 100194. https://doi.org/10.1016/j.caeai.2023.100194
[13] Else H. (2023). Abstracts written by ChatGPT fool scientists. Nature, 613(7944), 423. https://doi.org/10.1038/d41586-023-00056-7
[14] Essel, H. B., Vlachopoulos, D., Essuman, A. B., & Amankwa, J. O. (2024). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs). Computers and Education: Artificial Intelligence, 6, 100198. https://doi.org/10.1016/j.caeai.2023.100198
[15] Firat M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. Journal of applied learning and teaching, 6(1). https://doi.org/10.37074/jalt.2023.6.1.22
[16] García-Martínez, I., Batanero, J M F., Fernández-Cerero, J., & León, S P. (2023, January 15). Analysing the Impact of Artificial Intelligence and Computational Sciences on Student Performance: Systematic Review and Meta-analysis. Journal of New Approaches in Educational Research, 12(1), 171-171. https://doi.org/10.7821/naer.2023.1.1240
[17] Garrel, J V., & Mayer, J. (2023, November 9). Artificial Intelligence in studies—use of ChatGPT and AI-based tools among students in Germany. Humanities and Social Sciences Communications, 10(1). https://doi.org/10.1057/s41599-023-02304-7
[18] Gopalan, M., Rosinger, K., & Ahn, J. (2020). Use of Quasi-Experimental Research Designs in Education Research: Growth, Promise, and Challenges. Review of Research in Education, 44, 218-243.
[19] Grieve, R., Woodley, J., Hunt, S E., & McKay, A. (2021, July 15). Student fears of oral presentations and public speaking in higher education: a qualitative survey. Journal of Further and Higher Education, 45(9), 1281-1293. https://doi.org/10.1080/0309877x.2021.1948509
[20] Han, C., Feng, J., & Qi, H. (2024). Topic model for long document extractive summarization with sentence-level features and dynamic memory unit. Expert Systems with Applications, 238(Part B), 121873. https://doi.org/10.1016/j.eswa.2023.121873
[21] Liu, M., Zhang, L. J., & Biebricher, C. (2024). Investigating students' cognitive processes in generative AI-assisted digital multimodal composing and traditional writing. Computers & Education, 211, 104977. https://doi.org/10.1016/j.compedu.2023.104977

[22] Liu, Y., Chen, L., & Yao, Z. (2022). The application of artificial intelligence assistant to deep learning in teachers' teaching and students' learning processes. Frontiers in Psychology, 13.
[23] Luo, J. (Jess) (2024) A critical review of GenAI policies in higher education assessment: a call to reconsider the "originality" of students' work , Assessment & Evaluation in Higher Education, DOI: 10.1080/02602938.2024.2309963
[24] Miyazaki, K., Murayama, T., Uchiba, T. et al. Public perception of generative AI on Twitter: an empirical study based on occupation and usage. EPJ Data Sci. 13, 2 (2024). https://doi.org/10.1140/epjds/s13688-023-00445-y
[25] Mogavi, R H., Deng, C., Kim, J J., Zhou, P., Kwon, Y D., Metwally, A H S., Tlili, A., Bassanelli, S., Bucchiarone, A., Gujar, S., Nacke, L E., & Hui, P. (2024, January 1). ChatGPT in education: A blessing or a curse? A qualitative study exploring early adopters' utilization and perceptions. Computers in Human Behavior Artificial Humans, 2(1), 100027-100027. https://doi.org/10.1016/j.chbah.2023.100027
[26] Montebello, M. (2021). Personalized Learning Environments. 2021 International Symposium on Educational Technology (ISET), 134-138
[27] Moorhouse, B. L. (2024). Beginning and first-year language teachers' readiness for the generative AI age. Computers and Education: Artificial Intelligence, 6, 100201. https://doi.org/10.1016/j.caeai.2024.100201
[28] Nam, B H., & Bai, Q. (2023, November 16). ChatGPT and its ethical implications for STEM research and higher education: a media discourse analysis. International Journal of STEM Education, 10(1). https://doi.org/10.1186/s40594-023-00452-5
[29] Perrotta, C., & Selwyn, N. (2019). Deep learning goes to school: Toward a relational understanding of AI in education. Learning, Media and Technology, 45(3), 251-269.
[30] Schneider, B., Saw, G., & Broda, M. (2016). A Future for the National Education Longitudinal Program. AERA Open, 2.
[31] Stahl, B. C., & Eke, D. (2024). The ethics of ChatGPT – Exploring the ethical issues of an emerging technology. International Journal of Information Management, 74, 102700. https://doi.org/10.1016/j.ijinfomgt.2023.102700
[32] Vasconcelos, H., Jörke, M., Grunde-McLaughlin, M., Gerstenberg, T., Bernstein, M., & Krishna, R. (2022). Explanations Can Reduce Overreliance on AI Systems During Decision-Making. Proceedings of the ACM on Human-Computer Interaction, 7, 1-38.
[33] Wang, T., Diaz, D., Brown, C., & Chen, Y. (2023, October 3). Exploring the Role of AI Assistants in Computer Science Education: Methods, Implications, and Instructor Perspectives. https://doi.org/10.1109/vl-hcc57772.2023.00018
[34] Zhai, X., Chu, X., Chai, C S., Jong, M S., Starčič, A I., Spector, M., Liu, J., Jing, Y., & Li, Y. (2021, April 20). A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. Complexity, 2021, 1-18. https://doi.org/10.1155/2021/8812542
[35] Zhai, X., Nyaaba, M. & Ma, W. Can Generative AI and ChatGPT Outperform Humans on Cognitive-Demanding Problem-Solving Tasks in Science?. Sci & Educ (2024). https://doi.org/10.1007/s11191-024-00496-1