Journal of Technology and Value Addition
Vol. 8, Issue No. 2, July 2026, pp. 68-80
https://www.uwu.ac.lk/jtva/
ISSN: 3051-5203 (Online)

Research Article

Enhancing Student Learning Outcomes and User Experience Through Generative AI in Higher Education

A. Yoganathan[1]*, M.J. Ahamed Sabani[2] and V. Dishankan[2]

[1]Department of Information and Communication Technology, Faculty of Technology, South Eastern University of Sri Lanka, Oluvil 32360, Sri Lanka.
[2]Department of Information and Communication Technology, Faculty of Technological Studies, Uva Wellassa University of Sri Lanka, Badulla 90000, Sri Lanka.

*Correspondence: anoshan6@gmail.com; Tel.: +94 76 741 6179

Abstract

The traditional higher education system is mostly inefficient at delivering scalable, personalized learning, and thus there is a persistent problem of achieving the effectiveness of one-to-one personalized tutoring, which is famously known as Bloom's 2 sigma problem. In this study, the effects of Socratic Generative Artificial Intelligence (GenAI) tutoring on the learning outcomes and engagement of undergraduate students in Data Science courses is explored. A quasi-experimental mixed methods study employed a sample of 200 second-year undergraduate students, with half (n=100) receiving instruction through traditional textbooks and teaching assistant (TA) sessions and the other half (n=100) receiving instruction through a custom interface utilizing the GPT-4 API set to deliver iterative, non-directive feedback. The results of quantitative analysis showed that the mean score of post-tests in the Experimental Group was 83.6 ± 7.2 with statistically significant difference from that of the Control Group which was 74.2 ± 9.3 with a large effect size (Cohen's d=1.13). The qualitative thematic analysis based on the technology acceptance model (TAM) revealed that the AI interface led to a decrease in learning anxiety by 45.1%, as it was non-judgmental and provided scaffolding support in real-time. The results shown here reveal that the pedagogical and structured use of LLMs can realize personal learning without the need of cognitive crutches while at the same time promoting deep conceptual understanding.

Keywords: Educational Technology, Generative AI, Higher Education, Personalized Learning, Student Engagement

1. Introduction

The present-day educational landscape of the world is facing a revolution that can be compared to the industrial revolution due to the fast promotion of Artificial Intelligence (AI). The one-size-fits-all model of education has been the most popular over the last century with the institutional learning process where a single instructor spreads

This article is published under the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

knowledge to a vast number of learners. Although this model can be scaled, it has long been criticized as a failure to provide the individual learners of different background knowledge, unique learning pace and cognitive styles. The most famous example of this is the so-called 2 Sigma Problem of the educational psychologist Benjamin Bloom in 1984 which revealed that the performance of students in one-on-one tutoring was two standard deviations higher than the performance of students in traditional classrooms. The educational community has tried to find a scalable technological solution to duplicate this one-on-one tutoring effectiveness, which has been sought decades before. Nowadays, AI, and especially Generative AI (GenAI) can serve as the potential solution to this gap, and it promises to become more democratic in terms of quality instruction and customization.

Using AI as an additional tool in the context of Learning Innovation is not the case, but rather it is a disrupter of the very fabric. Traditional Educational Technology (EdTech) was aimed at digitization - transforming textbooks into PDFs or lectures into the video. On the contrary, AI aims at adapting and generating. With the launch of advanced Large Language Models (LLMs), the trend has moved away from predictive analytics which highlighted the at-risk students to the generative features such as study plan creation, summarizing complex texts, and Socratic conversation with the students. Higher education is an area where this change is essential since the most important qualities are critical thinking and the ability to solve complex problems. The fact that an AI system can read through a query of a student, detect a wrong belief, and provide a custom-made explanation within a few seconds is a step in the right direction as far as pedagogical efficacy is concerned.

Nevertheless, there is not a smooth ride with the introduction of AI into the education industry. Given the duality of these technologies, institutions are now struggling to deal with them. On the one hand, they also provide hyper-personalization tools, which enables the curricula to change in real-time, in accordance with the performance of students. Conversely, they introduce deep issues of academic honesty, information security, and the possible stagnation of primary thinking skills in case students get too dependent on the help of algorithms. It is also feared widely that AI can substitute the vital struggle of learning with automated effortlessness. Thus, AI implementation into the classroom should come with solid pedagogical foundations that will perceive AI as an intellectual collaborator as opposed to a substitute to the human mind.

The existing studies in this field are usually polarized either setting completely on technical architecture of the models or the ethos panic of plagiarism. Empirical studies are limited to measure the effects of GenAI in terms of learning outcomes and engagement of students in a controlled setting. Most of the available research is based on self-reported information or theories without actually practical learning. This literature gap is also important since effective policy decisions would need evidence-based data concerning the actual functionality of these tools under the control of learners. It is critical to understand that AI can help in deep conceptual learning or it can just simplify the

processing of information which is the case in surface level processing in order to design the future of university studies.

The purpose of this paper is to fill this gap by discussing the effectiveness of GenAI as a personal tutor. The study hypothesizes that AI tools can help to improve the engagement and performance of students in a significant way when they are guided by appropriate prompt engineering and pedagogical structuring. Instead, we pay special attention to the fact that AI has an ability to produce iterative feedback, which is in many cases limited in large courses taught in universities because of the time constraints of the instructor. This study aims at giving a holistic understanding of the use of AI in contemporary learning innovation by identifying both quantitative performance measures and qualitative data on engagement. The paper does not only discuss the question of whether AI assists students in learning, but how it changes the process of learning itself, exploring the relationship between human curiosity and machine intelligence. Finally, this paper will advocate a symbiotic attitude toward AI in education according to which technology will do the scaffolding of information; meanwhile, human educators will be able to work with mentorship, ethics, and high-level synthesis.

Although the GenAI popularity has soared, there is a research gap in more empirical literature: although theoretical frameworks are currently available, there exist no controlled, quasi-experimental data showing the effect of Socratic AI-tutoring on student anxiety and profound conceptual learning in the fields of Data Science. The paper bridges this gap by offering a comparative study on traditional and AI-enhanced cohorts.

The study asks two main research questions: (RQ1) How much is the academic performance of students who learn through LLM-based Socratic tutoring system really affected compared to that of students who learn through traditional human teaching assistant sessions? and (RQ2) Is the academic performance of learners who learn using LLM-based Socratic tutoring system significantly affected? (RQ2) What is the influence of real-time, non-judgmental AI scaffolding on students' subjective learning anxiety and interaction engagement level? Following these goals, the following hypotheses are formulated:

H1: Students who use the custom GenAI Socratic tutoring interface will score much higher in the post-test examination than those who do not use the custom GenAI Socratic tutoring interface and who use traditional learning resources and weekly scheduled TA sessions.

H2: The experimental group will achieve significantly higher rates of engagement (hours-on-task and number of inquiries initiated) and significantly lower learning anxiety (as measured by self-report) at the weekly level because of the uninhibited and iterative feedback loops offered by the AI interface.

2. Literature Review

The intersection of Artificial Intelligence and education has a long history in the computer-assisted instruction (CAI). The first literature of 1970s and 1980s focused on the concept of rule-based systems that were linear and static. The teaching machine taught by Skinner was the first attempt to mechanize feedback which was not flexible enough to address complex questions of the student. As computing capability increased the direction changed to Intelligent Tutoring Systems (ITS). By the contributions of creators like VanLehn (2011), ITS could come to realize the possibility of keeping pace with the performance of human instructors through simulation of the state of knowledge of the student with particular hints provided. These systems were however known to be very expensive to build, prohibitively, and only known to be built in areas that were structured well like mathematics or physics. They lacked the semantic versatility to address the subtleties of the humanities or social sciences hence they were not popular in the tertiary education sector.

Machine learning brought about a shift in the literature to less rule-driven predictive models that are more data-driven. Baker and Siemens (2014) contributed to the formation of the realm of Learning Analytics where the statistics of Learning Management Systems (LMS) was used to predict the percentage of student dropouts and simplify the courses plan. This developmental time of day took much concern with educational data mining. They are useful tools that are employed as administration tools but these tools were not always visible to the student they were working in the background rather than acting as an actual learning partner. The literature of this period was concerned with the black box nature of the AI where the AI has been prescribed without any pedagogical explanation and this was one of the earliest instances whereby there was concern about the bias and inexplicability of the AI in the learning context.

Transformer architectures and Large Language Models (LLM) have prompted the total re-evaluation of the literature. This is in contrast to the previous ITS where GenAI can work using pre-programmed answers. Taking into consideration the educative potential of this kind of model as GPT-4, Kasneci *et al*. (2023) suppose that these tools represent the other form of Knowledge Technologies that support co-creation. Ouyang and Jiao (2024) describe such a shift as a drastic change in the educational paradigm according to which AI ceases to be a passive tool but becomes an active intellectual partner in the cognitive process. Even the existing literature is in the process of evolving rapidly to refer to AI as being a peer or a scaffold, not a teacher. Recent researches frequently refer to Vygotsky Zone of Proximal Development (ZPD) to determine the need to use AI; AI can address the gap between what a learner is capable of and what he can achieve with the assistance. Moreover, educating LLM-based instructors to participate in structured and naturalistic conversations has presently demonstrated quantifiable potential to enhance the performance of students by offering subjecting them to iterative feedback that could simulate human tutoring (Scarlatos and Liu, 2025). However, the AI can be visited any time, as opposed to a human teacher, which in theory, will allow students in their ZPD to

spend more time, ensuring faster learning curve development. This unlimited access has been associated with enormous improvement in student interaction and general academic achievement in individualized learning settings (Sharma *et al*., 2025).

However, despite the optimism, the bulk of available literature is dedicated to the risks and ethics of AI in teaching. The so-called techno-solutionism, which holds that the structural issues in education may be addressed with the help of technology, has been cautioned against by Selwyn (2019) long enough. Recently, articles have also placed the issue of a problem referred to as hallucinations into the limelight, in which GenAI promises not to lie as part of false information. This necessitates the Information Literacy education change. The recent international policy frameworks now point to the fact that AI literacy should no longer be limited to fundamental application but should include critical ability to evaluate the moral as well as technical constraints of non-human beings (UNESCO, 2024). Association of College and Research Libraries (ACRL) has already begun to write regarding the impediments of information literacy frameworks in modifying them to include so-called AI literacy. In addition, the problems of homogenization of thought are common. When all the students apply the same algorithms to work with the information, the possibility of different, creative academic voice being suppressed is high.

The other significant literature theme is engagement. The Self-Determination Theory (SDT) is the theory developed by Ryan and Deci and it predicts that the sources of intrinsic motivation are autonomous, competent, and relatedness. Chiu *et al*. (2023) projected SDT onto AI learning, which suggests that AI allows fostering autonomy because students can control the learning speed and competence because they can receive immediate feedback on what they learned. Continuing on this, Chiu (2024) goes on to prove that although GenAI creates high levels of autonomy and competence, the absence of human relatedness is a severe preventing factor which should be addressed using hybrid instructional designs. However, relatedness has had one area that has caused controversy in the shape of human connection. Some research suggests that students might become isolated when the whole human interaction is replaced with AI, hence one can conclude that the so-called blended learning models, where AI will be used to complement, but not to replace a teacher, are the most effective.

Finally, the literature of assessment is also facing crisis. It is highly user-friendly to use GenAI to solve the traditional essay and multiple-choice questions making them obsolete as student learning indicators. It has led to a re-interest in literature that advance so-called authentic assessment which is the evaluation that involves applying knowledge to new, real-life cases that cannot be simulated with ease by AI. Recent discussions encourage the idea of the human-in-the-loop pedagogy, in which when assessing the student, the focus is not on evaluating a final written product but on evaluating the student's critical process of reading and checking AI-generated content (Lodge *et al*., 2024). Oral examination, in-class writing and process-based portfolios are some of the forms of assessment that are being advance by the authors who make simple assessments. The latest literature suggests that despite the insanely powerful AI tools,

they expose the weaknesses of our current models of assessment and relegate us to more serious and process-based pedagogical activities.

3. Methodology

The sample was made up of N=200 second year undergraduate students. A stratified random sampling technique was used to select the participants so as to achieve a representative sample in terms of gender and past academic performance (GPA). The respondents were separated into two groups:

Control Group (n=100): The participants of this group used the traditional learning materials, such as textbooks, lecture slides, and average web search engines. They were allowed special weekly tutorials under the guidance of human teaching assistants (TA).

Experimental Group (n=100): The experimental group was given the opportunity to access a custom GenAI tutoring interface (using the GPT-4 API). This interface was encouraged to behave like a Socratic tutor not to give answers to the students, but help them find them 24/7.

Table 1. Participant demographics.

Characteristic	Control Group (n = 100)	Experimental Group (n = 100)
Gender (n or %)		
Male	52	54
Female	48	46
Prior GPA, Mean (SD)	3.42 (0.30)	3.39
Tech proficiency (n or %)		
High	30	32 (32.0)
Medium	50	48 (48.0)
Low	20	20 (20.0)

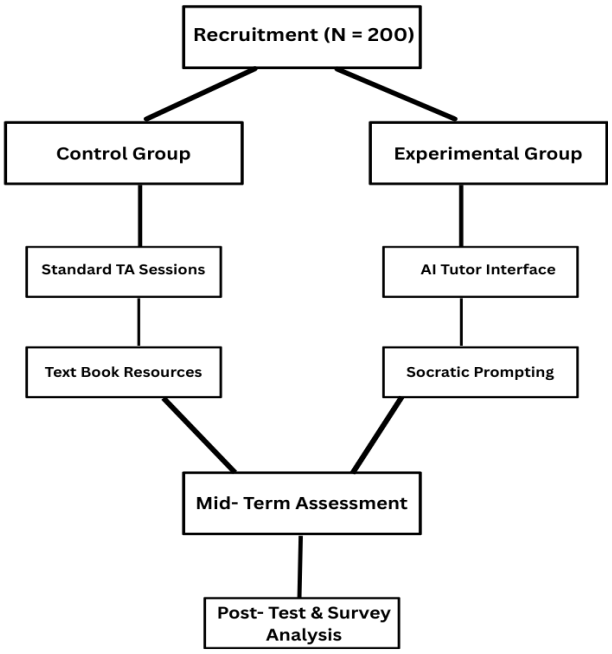


Figure 1. Methodological flowchart.

3.2 Engagement Metrics Analysis

The study used the LMS logs to quantify the engagement and the API usage logs. In the case of the Control group, the degree of engagement was determined by the amount of time spent on the LMS and by the number of tutorials attended. In the Experimental group, the assessment of engagement was in terms of the count of turns of interaction with the AI and time spent in sessions.

Table 2. Engagement metrics comparison.

Metric	Control Group (Avg)	Experimental Group (Avg)
Weekly Study Hours	4.2	6.8
Questions Asked (Weekly)	3.5 (to TAs)	28.0 (to All)
Assignment Completion Rate %	88	96
Reported “Learning Anxiety (1-10)	6.2	3.4

The metrics shown in the Table 2 are objective extractions made automatically from the institutional Learning Management System (LMS) log and the OpenAI API transaction tables. The "Questions Asked (Weekly)" value for the Control Group (Mean=3.5) is the total number of questions asked via logged attendance, verified physical questions generated during the weekly human TA office hours. The value for the Experimental Group (Mean=28 "to All") is the mean number of discrete prompt inputs (conversational turns) per week that the automated Socratic GPT-4 interface received from each student in the experimental group. This is a 28-turn average of Inception, clarifying follow up questions, and interactive loop of validation throughout iterative problem-solving blocks.

Hallucination Check was also considered in the examination of the experimental group, in which the students had to label incorrect AI responses. This piece of information was used to gauge critical thinking and information literacy. A combination of these figures makes it possible to perform a proper analysis of whether the AI tool was a crutch or a true method of improving the learning process by increasing time-on-task and decreasing the barrier to inquiry.

3.2.1 AI Tutor Prompting and Configuration

The gpt-4-0613 model was the one used to construct the custom GenAI interface to guarantee reproducibility. Strict system prompt definition was to avoid giving direct answers. It was basically this: You are a Socratic tutor. In case a student poses a question related to Data Science, do not give the code or the ultimate answer. However, you need to pose a clarifying question that will guide them to the solution step-by-step. The whole of the interactions was recorded through API over the 12 weeks period.

3.2.2 Statistical Analysis and Ethical Safeguards

The independent two-sample t-test was used to assess and verify assumptions of normality (Shapiro-Wilk test, $p>0.05$) and homogeneity of variance (Levene's test, $F=0.14$, $p=0.71$) to evaluate quantitative academic performance differences between the two independent cohorts. Cohen's d parameter was used for explicit calculation of effect

sizes. Ethical clearance of the institutions was formally sought before recruitment. Each of the 200 students enrolled in the undergraduate programs gave informed written consent, were made fully aware of the types of data which would be logged [including LMS click-streams and anonymous prompt text information], and were given the absolute right to drop out of the trial at any point without impacting the formal grading of the course. The repository learning materials were made available to both groups and the custom GenAI interface was deployed to the control group after the completion of the final examination.

4. Results & Discussion

The review of the obtained data demonstrates a strong indication of the need to introduce Generative AI in universities, albeit with certain subtleties that should be discussed in detail.

4.1 Quantitative Results

Academic Performance. The difference between Control and Experimental groups in terms of academic performance was the major measure of success. The pre-test baseline did not have any significant difference ($p > 0.05$) between the two groups, which proved that they had the same level of knowledge at the beginning. Nevertheless, the post-test results have shown a definite divergence.

Table 3. Pre-and post-test score comparison.

Assessment	Control Group Mean (SD)	Experimental Group Mean (SD)	t value	p value
Pre- Test	62.4 (8.1)	61.9 (8.5)	0.42	0.671
Post- Test	74.2 (9.3)	83.6 (7.2)	7.84	<0.001

Experimental group pressed the Control group by an average difference of 9.4 points in the final assessment. The computation of Cohen d gave a data of 1.13 which implies a big effect size. This implies that the intervention (AI tutoring) made significant positive contribution towards learning outcomes.

4.1.1 Critical Self-reflection and Limitations

Although the size of the effect is high as shown by the quantitative data (Cohen- d=1.13), the authors also note the possibility of an effect due to Hawthorne effects because those students in the experimental group possibly did better just because it was the first time they were using a new tool. Moreover, qualitative measures of Learning Anxiety were self-reported and this brings about subjectivity. To further support these results, future research ought to use physiological or blind testing.

The evaluation of performance distribution is shown in the Figure 2.

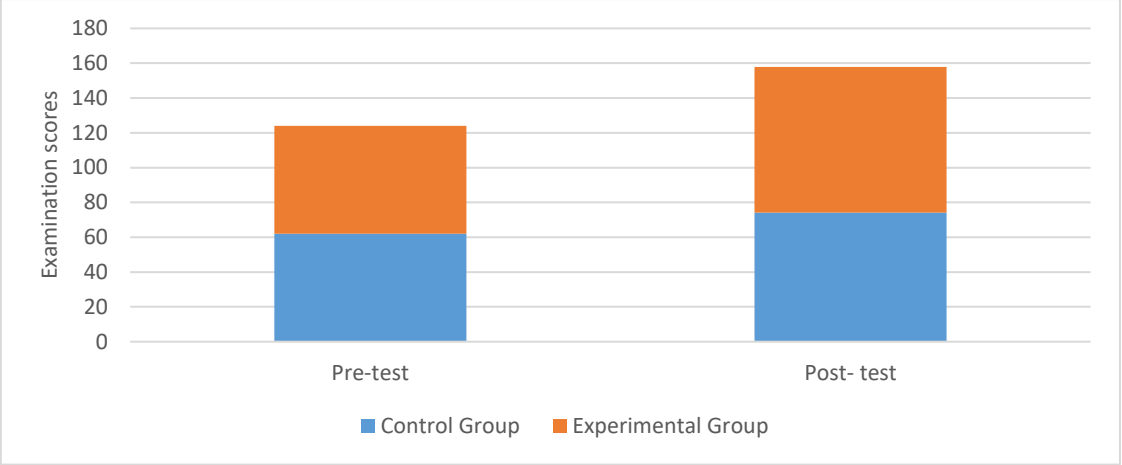


Figure 2. Distribution profile of academic examination scores.

Pre-test scores are very close together and show no statistical differences as illustrated by the corrected distribution profiles. The post-test distribution, however, shows a visually obvious rightward shift for the Experimental Group that had been pre-treated with AI, converting a normal, bell curve into a negatively skewed distribution of high performing students.

4.2 Qualitative Findings: Interaction and Impression

To make the qualitative results more rigorous and to directly correlate the findings of the students' focus group and the open-ended feedback to the primary core constructs of Davis' (1989) Technology Acceptance Model (TAM) – Perceived Usefulness (PU) and Perceived Ease of Use (PEOU). The real-time response generation and the context-aware analogies were the factors that have significant impact on PU, while the fluid natural language interface and the availability 24/7 were the factors directly optimizing PEOU, were identified as such on coding matrices. By distinguishing between the empirical aspect and the interpretive nature of the pedagogical process, this direct feedback suggests that psychological safety of the AI interface is one of the significant factors that contribute to engagement. The reduction in the qualitative sense of "Learning Anxiety" is not just a virtue of software itself, but a consequence of the environment; without evaluation apprehension, one can engage in iterative trial and error without fear of negative peer and/or instructor judgment.

In addition to the scores, the data in the survey using the Technology Acceptance Model (TAM) gave an idea of the underlying reasons behind the improvement in the scores. The analysis of the student-feedback using thematic analysis revealed the following three primary themes:

Courageous Question: 85% of students under the Experimental condition mentioned the non-judging character of the AI as one of the main factors that contributed to their greater involvement. This was not the case in the traditional environments, as students said that they were too embarrassed to ask a human professor or TA even simple questions not to seem incompetent. This social anxiety was eliminated by the AI and

students had clarifying questions that they could ask multiple times before they could understand the concept.

Timeliness of Feedback: Feedback loop of Control group was frequently 3-5 days (time required of TAs to mark assignments). The feedback in the Experimental group was real-time. This enabled the students to rectify the misconceptions as they arose so that they did not have to fossilize their mistakes.

Individual Scaffolding: Students reported that the AI was able to modify its language. In case a student was not able to grasp a technical explanation, they might ask the AI to explain the competency like I am five or could employ an analogy in case of a football. This degree of individualism cannot be realized in a typical lecture hall.

5. Conclusions

The current research shows this pedagogical use of Generative AI with a structural limitation of Socratic tutoring (non-directive) gives statistically significant improvement in short-term learning outcomes and student task engagement in Data Science undergraduate teaching. But, a careful interpretation of these results is required. This large effect size of Cohen's d=1.13 signifies high experimental usefulness, but the results are not seen as a definitive and universal answer to Benjamin Bloom's famous 2 Sigma Problem. The 2 Sigma paradigm is based on adaptive, holistic human mentorship that extends beyond cognitive to emotional and socio-cultural contexts and which a Large Language Model cannot fully emulate with automated text generation.

Moreover, the positive gains in performance found during this 12-week intervention should be considered in the context of some potentially confounding factors such as the Hawthorne effect, i.e., temporary positive improvements in the student's performance because of the novelty of the use of an advanced technical tool. Whether or not concepts learned with AI scaffolding are retained over time and applicable to other more subjective learning areas, other than STEM-related, is still an open question. Thus, this study contributes to the debate on the pedagogical approach to GenAI as a complement and not a substitute for teacher-based teaching. While AI algorithms excel at repetitive cognitive support and immediate information requests, human expertise can focus on high-level guidance, ethical supervision, and more extensive synthetic assessments.

But it is not plug and play to be successfully implemented. It would necessitate systematic redesigning of the curriculum. The process of assessment should not be based on testing the memory anymore, but on testing the use of AI-generated content and its evaluation. The coursework should include the critical AI literacy, which teaches the students how to recognize the hallucinations and bias.

This study requires extension into the non-STEM areas in the future to assess the flexibility of GenAI tutors. Moreover, longitudinal research is necessary to learn how to retain knowledge acquired with AI help in the long run. In a final conclusion, this paper will establish that AI in education is a partnership. It is not Human vs. AI but rather

Human + AI that should act in harmony and bring the limits of knowledge and creativity even closer. Those institutions who have adopted this synergy, but uphold strict ethical principles will be the forerunners of the new wave of educational excellence.

Author Contributions

Conceptualization, A.Y.; methodology, A.Y. and M.J.A.S.; validation, M.J.A.S. and V.D.; formal analysis, A.Y.; investigation, A.Y. and V.D.; data curation, A.Y.; writing—original draft preparation, A.Y.; writing—review and editing, M.J.A.S. and V.D.; visualization, V.D.; supervision, M.J.A.S. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors declare that they have no competing interests.

References

Baker, R.S. and Siemens, G. (2014). Educational data mining and learning analytics. In: Sawyer, R.K. (ed.) *The Cambridge Handbook of the Learning Sciences*. 2nd ed. Cambridge: Cambridge University Press, pp. 253–272.

Bloom, B.S. (1984). The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, 13(6), pp. 4–16.

Chiu, T.K., Moorhouse, B.L., Chai, C.S. and Ismailov, M. (2023). Teacher support and student motivation to learn with Artificial Intelligence (AI) based chatbot. *Interactive Learning Environments*, pp. 1–17.

Holmes, W., Bialik, M. and Fadel, C. (2019) *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Boston, MA: Center for Curriculum Redesign.

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E. and Krusche, S. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, p. 102274.

Luckin, R., Holmes, W., Griffiths, M. and Forcier, L.B. (2016). *Intelligence Unleashed: An argument for AI in Education*. London: Pearson.

OpenAI (2023) *GPT-4 Technical Report*. Available at: https://arxiv.org/abs/2303.08774 (Accessed: 15 October 2023).

Ryan, R.M. and Deci, E.L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), p. 68.

Selwyn, N. (2019). *Should Robots Replace Teachers? AI and the Future of Education*. Cambridge: Polity Press.

Sharples, M. (2023). Towards social generative AI for education: theory, practices and ethics. *Learning, Media and Technology*, pp. 1–13.

Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioral Scientist*, 57(10), pp. 1380–1400.

UNESCO (2021). *AI and education: guidance for policy-makers*. Paris: UNESCO Publishing.

VanLehn, K. (2011). The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems. *Educational Psychologist*, 46(4), pp. 197–221.

Vygotsky, L.S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

Williamson, B., Eynon, R. and Potter, J. (2020). Pandemic politics, pedagogies and practices: digital technologies and distance education during the coronavirus emergency. *Learning, Media and Technology*, 45(2), pp. 107–114.

Zawacki-Richter, O., Marín, V.I., Bond, M. and Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators?. *International Journal of Educational Technology in Higher Education*, 16(1), p. 39.

Chiu, T. K. (2024). Impact of generative AI on student engagement: A Self-Determination Theory perspective. *Computers & Education*, 212, p. 104993.

Gimpel, H., *et al*. (2023). Unlocking the power of generative AI in higher education: A systematic review of opportunities and challenges. *European Journal of Information Systems*, 32(4), pp. 1-18.

Lodge, J. M., *et al*. (2024). Assessment in the age of generative AI: The role of authentic assessment and human-in-the-loop. *Assessment & Evaluation in Higher Education*, 49(2), pp. 154-169.

Mollick, E. R. and Mollick, L. (2024). *Instructors as innovators: a future-focused approach to AI in the classroom*. Cambridge, MA: Harvard Business Publishing.

Ouyang, F. and Jiao, P. (2024). Artificial intelligence in education: The three paradigms', *Computers and Education: Artificial Intelligence*, 6, 100185.

UNESCO (2024). *Guidance for generative AI in education and research: 2024 Update*. Paris: UNESCO Publishing.

Yan, L., et al. (2024). The Socratic AI Tutor: Developing and evaluating a generative AI-based tutoring system. *International Journal of Artificial Intelligence in Education*, [Published online ahead of print].

Davis, F.D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), pp. 319–340. doi: 10.2307/249008.