ISBN: 978-628-95207-4-3. ISSN: 2414-6390. Digital Object Identifier: 10.18687/LACCEI2023.1.1.1504

Artificial intelligence tutoring versus tutoring with experts in learning the preclinical and clinical areas of medicine

José Augusto Rojas Peñafiel, MD[1], Ana Gabriela Valladares Patiño, MSc[2]
[1]Health Sciences Faculty, Universidad Internacional SEK (UISEK), Equator, jose.rojas@uisek.edu.ec,
[2]Universidad Central del Ecuador, Equator, agvalladares@uce.edu.ec

Abstract– The aim of this study was to compare the effectiveness of AI-based tutoring with expert tutoring in Pharmacology for preclinical and clinical areas among 100 randomly selected sixth and tenth-semester university students. Both types of tutoring were provided for a duration of six weeks. A baseline was established to ensure a comparative analysis between the two tutoring methods. The hypothesis was verified, as students who received AI-based tutoring demonstrated better performance in the multi-test examination at the end of the intervention period. Therefore, the findings of this study provide valuable insights for the development of medical curricula at universities, considering AI as a tool to enhance teaching and learning processes.

It is important to note that both groups received expert-led classes in the same manner, and this study solely modified the process of extra-class tutoring for the Pharmacology subject. Furthermore, the study evaluated the same topics and subtopics covered in the tutorial sessions.

Keywords-- Artificial Intelligence, expert tutoring, tutoring, Medical Education, Effectiveness.

I. Introduction

Medical education is a complex and demanding process that requires students to acquire a vast amount of knowledge and skills in both preclinical and clinical areas [1]. In recent years, artificial intelligence (AI) has been proposed as a potential solution to enhance medical education outcomes. One application of AI in medical education is the use of intelligent tutoring systems, which provide personalized feedback and guidance to individual students [2]. The objective of this study is to explore the use of AI tutoring systems in learning preclinical and clinical areas of medicine, specifically in the field of pharmacology.

The integration of intelligent tutoring systems in medical education offers several advantages [3]. These systems enable personalized tutorials where the system can assess the student's level of knowledge and identify areas that require further reinforcement [4]. As students engage in the suggested activities, the difficulty level can be adjusted, and guidance can be provided based on their strengths and weaknesses. These systems are integrated into learning management systems, which have experienced significant growth.

Intelligent tutorial systems that are integrated with virtual reality simulation allow students to experience hyper-realistic tool interactions with anatomical simulators [5]. These systems can leverage large datasets and employ machine learning to differentiate experiences and provide feedback on operational performance [6]. By breaking down teachable psychometric skills, students can receive personalized feedback on specific factors identified by algorithms, leading to improved performance. Virtual reality simulation and machine learning algorithms can objectively quantify performance and enhance the accuracy of essential medical knowledge concepts [7]. Moreover, these systems can empower medical educators to develop more qualitative and quantitative training with summative assessment tools to meet future challenging pedagogical requirements [8].

The topic of artificial intelligence (AI) in medical education and training is increasingly relevant in research, as it holds numerous potential benefits. Specifically, comparing AI mentoring with expert-led mentoring in learning preclinical and clinical areas of medicine is of great interest. AI tutoring systems have the potential to reduce skill heterogeneity and complement competency-based curricular training. Simultaneously, the utilization of virtual reality simulation and machine learning algorithms can objectively quantify performance and improve the accuracy of key medical knowledge concepts. Therefore, it is crucial to evaluate the effectiveness of AI mentoring systems compared to traditional approaches in medical education and training [9].

However, the effectiveness of AI mentoring compared to traditional expert instruction in medical education remains unclear. Hence, this study aims to compare the outcomes of AI tutoring versus expert tutoring in learning preclinical and clinical areas of medicine, with a specific focus on pharmacology. By understanding the relative effectiveness of these two methods, potential areas for improvement in medical education can be identified, ultimately enhancing the quality of care provided by future healthcare professionals.

Moreover, it is yet unknown whether the availability of AI systems 24 hours a day, seven days a week, 365 days a year influences performance improvement compared to the hours established and limited by tutorials provided by expert teachers.

The aim of this study is to compare the effectiveness of AI-based tutoring with that delivered by experts in teaching the preclinical and clinical areas of medicine. To achieve this goal, a randomized controlled study will be conducted, where participants will be randomly assigned to one of two groups: the AI-based mentoring group and the expert-led mentoring group. The hypothesis of this study is that AI-based tutoring will be more effective than expert-led tutoring in improving learning in the preclinical and clinical areas of medicine. To test this hypothesis, the learning outcomes of the two groups will be measured and compared.

This study could provide valuable information on the effective utilization of AI in medical education. If AI-based tutoring is demonstrated to be more effective than expert-led tutoring, it could have significant implications for medical education globally. For instance, AI-based tutoring could help address challenges faced by medical education, such as a shortage of expert teachers, the need for personalized learning experiences, and the necessity to keep pace with rapid advances in medical research. Moreover, this study could pave the way for the development of more effective AI-based tutoring systems that can enhance medical education worldwide.

II. METHODOLOGY

A. Participants

Undergraduate medical students were recruited via email, providing instructions on planning, timing, and the advantages/disadvantages of voluntary participation in this experiment. They were informed that participation was voluntary and that they could opt out at any time without any consequences. All data were processed confidentially. All participants signed an informed consent form, and the experiment was conducted in accordance with legal regulations. At the end of the experiment, participants received feedback on their results to improve their performance. The participants consisted of 100 sixth- and tenth-year medical students. They were randomly assigned to two equally sized groups. We used the block criterion for randomization, creating 10 blocks with 10 participants per group, ensuring an equal number of participants in both groups and an equal probability of assignment. Out of the 100 participants, all attended the follow-up session and were included in the analysis. Among the 100 participants, 50 were assigned to the expert tutor group, and 50 were assigned to the artificial intelligence tutor group. Additionally, out of the 100 participants, 59 were women (aged 21-26; M = 22.36, SD = 2.02), and 41 were men (aged 23-28 years; M = 22.27, SD = 2.65).

The group of expert tutors received training from an expert regarding the tests administered to the participants of the corresponding study group. The expectation was that the expert's comments would help participants focus on relevant aspects of task learning and integrate their declarative medical knowledge with their perceptual-cognitive skills. The artificial intelligence tutor group received immediate feedback at the end of each task or problem- or question-based learning. The expectation was that the feedback provided by the AI through a simulator would help participants acquire necessary perceptual-cognitive skills. The feedback from the AI tutor provides information about the accuracy of the acquired knowledge, helping students identify their strengths and weaknesses in medical concepts compared to their preparation group for each learning challenge.

Furthermore, the artificial intelligence tutor reminds participants how to assess themselves and provides insights into the clinical concepts behind the challenges, aiding in the recall of correct declarative knowledge. The feedback from the expert tutors guides participants' attention toward relevant aspects of the tasks and assists in integrating their knowledge with cognitive and perceptual skills.

B. Virtual Learning Experience Platform

The Virtual Learning Experience platform with artificial intelligence was used in this study and installed on the participants' computers.

C. Procedure and Design

Data collection for measurements in this experiment was conducted immediately after a training session and continuously based on the learning needs established by the study population, with a final retention test after 6 weeks. The training session lasted 2-3 hours. At the beginning of the experiment, participants received written information about the content they would learn at predetermined times, according to the baseline established with an initial standardized test for all participants. Subsequently, participants watched a video on how to use the AI-enabled platform, which provided instructions on utilizing the AI-powered learning tools it contained. To ensure that participants acquired all the necessary knowledge to use the platform, they were required to take a knowledge test on the computer and answer all the questions correctly. In case of any mistakes, participants received feedback on their answers and had to retake the test until they could answer all the questions correctly on two consecutive administrations of the platform-only test. The group of expert tutors received classes and evaluations from experts in each of the study areas over a 6-week period, following the schedule established by the baseline.

All participants were instructed to achieve the highest possible grades within a time restriction during assessment tests, with 40 questions per hour. After practicing on the test simulator, participants took a practice test. If an answer was incorrect, participants could practice that specific question or topic concept and repeat the test for that concept until they could answer correctly twice in a row. During the testing phase, all help screens in the simulator were disabled. At the end of the session, participants completed a short questionnaire evaluating the test simulator and the training they received.

After 6 weeks, participants were invited to return for the final test. They took the same knowledge tests as in the training session and completed questionnaires at the end. Ten experienced tutors with expertise in preparing standardized multi-test tests participated in this study. They provided feedback similar to what was done for the AI tutor groups.

D. Variable Results: Knowledge Test

The knowledge test consisted of 160 multiple-choice questions, divided into two blocks: preclinical and clinical, based on the participants' semester of study. The test was conducted on a computer without access to external reference materials, with a time limit of 4 hours. At the end of the test, all questions were automatically scored as either correct or incorrect. Participants received feedback for each question and for each block of the test, identifying learning themes and subthemes. The quality of each question was independently rated by six medical experts with mentoring experience. The experts, who were blinded to the groups and participants, assigned a rating from 1 to 4 points based on the following criteria:

Low complexity
Moderate complexity
High complexity
Very high complexity

E. Performance Measures

In addition to evaluating the quality and complexity of the questions, the time taken to complete each question during the computerized test was measured using a stopwatch. The number of attempts made by participants for each block of the knowledge test was also recorded.

F. Questionnaire

The questionnaire was completed immediately after the final testing session. It consisted of three parts. The first part collected demographic information about the participants and inquired about the frequency of their computer-based standardized test-taking experiences.

The second part consisted of nine questions regarding the instructions, the training experience, and the cognitive load, which were rated on a Likert-type scale ranging from 1 to 5.

The last part contained open-ended questions where participants could provide further feedback on the quality of the training and suggest improvements for the training session.

G. Data Analysis

All analyses were conducted using SPSS 21.0. The significance level (alpha) for determining statistical significance was set at 0.05. To investigate the influence of the two types of feedback on participants' acquisition of cognitive skills, the Kruskal-Wallis H test was performed, using the group as an independent variable and the number of attempts made by participants for the knowledge test, the number of attempts made by participants during test practice, and the time taken by participants to complete the test after the training sessions as dependent variables. Post hoc multiple comparison analyses were conducted to determine which groups differed based on the Kruskal-Wallis test.

III. RESULTS

Baseline Measurement

Two reference variables that could have influenced the experiment were whether participants had previously taken a timed standardized test and whether they had undergone multitest training. However, only two participants had taken an exam before, making it impossible to conduct an analysis of variance to determine group differences.

The second variable, the number of times participants had taken a standardized test, was not normally distributed. Therefore, we performed the Kruskal-Wallis H test as a nonparametric analysis of variance. The results showed no significant difference between the two experimental groups ($p=0.9$), indicating that all groups had comparable prior experience with multitest testing.

Training Session

During the training session, participants practiced until they achieved the competency level established in the six-week program. As a result, all participants attained the highest scores in knowledge.

Retention Session

The scores for image quality in the retention session are presented in Figure 1. An analysis of covariance for image quality revealed a significant difference between the groups. Subsequent multiple comparison analysis indicated that the AI group obtained significantly higher scores on the practical test compared to the expert group. There was no significant group difference in the time taken by participants to complete the final practice test.

Results

p-value and Statistical Significance

The two-tailed p-value is less than 0.0001, indicating an extremely significant difference from a statistical perspective according to conventional criteria.

Confidence Interval:

The mean difference between AI and Expert is 16.86.

The 95% confidence interval for this difference ranges from 12.88 to 20.84.

Intermediate Values Used in the Calculation:

t=8.4046

gl=98

standard error of difference=2.006

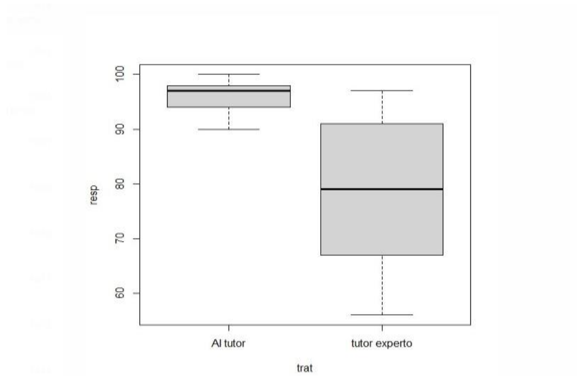
TABLE I
RESULTS OF AI TUTORING VS. EXPERT TUTORING

DATA		
Group	AI [a]	Expert
Mean	96,04	79,18
SD [b]	2,85	13,9
SEM [c]	0,4	1,97
n [d]	50	50

Artificial Intelligence [a]; Standard Deviation [b]; Structural equations model [c]; Sample size [d]

Figure l Results of AI tutoring vs. expert tutoring

Score



Evaluations

Fig. 1 Score was significantly higher in the AI tutoring group (p=0.0001).

Questionnaire

Overall, the participants found the instructions used in the training session to be very clear (median score of 4). They perceived the theoretical part of the training as easier compared to the practical part (median scores of 3 and 4, respectively) and indicated that they put a lot of effort into both parts (median score of 4 for both parts).

The video and the knowledge test were considered the most useful materials. Participants stated that both helped them in remembering and applying what they had studied.

IV. DISCUSSION

In this study, our aim was to determine the optimal way of guiding students to acquire and retain complex medical knowledge, with a focus on feedback sources.

Digital Object Identifier: (only for full papers, inserted by LACCEI).
ISSN, ISBN: (to be inserted by LACCEI).
DO NOT REMOVE

The participants in the study received feedback from two different sources: an expert tutor (Ex) and an AI Platform tutor (IA TUTOR). The feedback source was manipulated to compare the effects on the participants' learning outcomes.

During the baseline test, all participants started without any prior tutoring, making it difficult to compare their results at that point. However, during the final questionnaire, several differences between the groups emerged. The group of students with an expert tutor did not show a significant advantage in terms of response time, which is in line with our hypothesis.

In the artificial intelligence group, feedback effectiveness for medical education, based on question-based learning and clinical vignettes, showed a moderate advantage over expert feedback. The AI feedback resulted in reduced response time and improved performance.

These findings can be attributed, in part, to variations in the type and timing of personalized learning dedication for each student, focusing on different topics and subtopics of the pharmacology program undertaken over the course of 6 weeks, which influenced the development of cognitive skills tasks.

During the 6-week retention test, we observed differences in the time required to acquire conceptual knowledge for each topic and subtopic between the groups. In terms of learning quality, both groups showed improvement in their knowledge, but the group that received AI tutoring, which utilized individualized spaced repetition, achieved higher performance scores.

An important implication of this finding was that teaching and learning strategies should be tailored to the specific needs of different groups, particularly in the case of clinical cases based on image reading. The participants who were tutored by AI acquired more skills and demonstrated better performance in this type of question, as reflected in their retention exam with more correct answers in such tasks. [10]

The experts likely guided the students in acquiring skills such as the use and interpretation of clinical concepts, thereby reducing the metacognitive load associated with the use of image screens. [11-12]

Another significant fact was that the skills acquired by each group differed. However, when analyzed holistically, the use of spaced repetition and longer preparation time during the hours of student availability resulted in better outcomes for the groups tutored by AI. Among the two groups, the AI group demonstrated better consolidation of key concepts. This could be attributed to a reduction in cognitive load during the expert tutor sessions, which minimized the metacognitive load associated with relying extensively on help platforms, compared to the AI group. [13-20]

While simple skills benefit from terminal feedback at the end of a session, complex cognitive skills related to knowledge acquisition benefit from feedback provided during the training session, especially in the case of individual research or multitest question-based learning approaches. [21-26]

Therefore, by utilizing artificial intelligence, it is possible to divide the components of cognitive skills into declarative aspects with varying complexity based on each student's baseline level. This personalized study approach, conducted on a topic and subtopic basis, helps address the individual learning needs of students rather than relying solely on group-based study [27-32]. It leads to an improvement in individual and group performance by establishing required knowledge as learning objectives. Furthermore, it can assist in designing optimal feedback strategies for cognitive skill training by aligning teaching strategies with the specific type of knowledge demanded by each skill, particularly for case-solving and multitest questions.

In summary, while expert feedback may reduce the training time required to acquire a medical cognitive skill, it may not be optimal for long-term skill retention. AI feedback appears to promote better retention of complex medical skills among students [33-36]. However, whether this finding remains consistent after more practice sessions or longer retention intervals requires further investigation.

The integration of AI in this context is supported by our finding of a positive correlation between participant scores on standardized declarative tests, suggesting that this declarative knowledge is solidified through AI-prepared spaced repetition for each participant, resulting in better consolidation of key concepts [37-39].

In practice, cognitive skills training often emphasizes acquisition rather than retention. However, since the acquisition and retention of cognitive skills seem to be influenced by different feedback sources, it reaffirms the established notion that students vary considerably in the time required to acquire and sustain cognitive abilities, particularly in the medical field where these skills must be maintained over several years.

V. CONCLUSIONS

In conclusion, the use of AI in medical education holds great promise for enhancing learning effectiveness in the field. It has been established that AI tutoring leads to improved performance in multi-test assessments, indicating that AI tutoring is more effective than expert-led tutoring in teaching clinical and preclinical aspects of medicine. The potential implications of this study are significant and could potentially revolutionize medical education worldwide.

The results demonstrate a greater impact of AI in the study and analysis of images compared to expert-led tutorials, suggesting that repetitive exposure to similar types of questions and clinical cases, particularly those involving image interpretation, enhances image interpretation skills.

Tutorials conducted by AI required higher levels of interaction and dedication from students compared to face-to-face tutorials with an expert tutor. Therefore, it can be concluded that the time investment in AI tutorials surpasses that of traditional tutorials with an expert.Furthermore, by correlating the interaction time with the AI platform and the number of correct answers, it can be inferred that greater interaction and practice result in higher final grades for students. Conversely, students in the AI group who displayed inactivity in using the platform did not demonstrate improvement and even experienced a decline in performance compared to their baseline examination. This suggests that this group of students did not engage in any form of tutorial. However, this conclusion cannot be made for the expert-led tutoring group, as the time spent was the same for all students in that group.

It is important to note that AI tutorials provide greater evidence for educational audits, both internal and external. This aspect is crucial for decision-making by the authorities of higher education institutions.

Expressions of gratitude:

The authors would like to express their gratitude to F. Idrovo from LB-FCQ / UCE for his collaboration, as well as to the students who participated in the study.

REFERENCES

[1] McArthur D, Lewis M, Bishary M. The roles of artificial intelligence in education: current progress and future prospects. Journal of Educational Technology. 2005; 1(4):42–80.

[2] Schmidhuber J. Deep learning in neural networks: An overview. Neural networks. 2015; 61:85–117. https://doi.org/10.1016/j.neunet.2014.09.003 PMID: 25462637.

[3] Grapov D, Fahrmann J, Wanichthanarak K, Khoomrung S. Rise of deep learning for genomic, proteomic,and metabolomic data integration in precision medicine. Omics: a journal of integrative biology.2018; 22(10):630– 6. https://doi.org/10.1089/omi.2018.0097 PMID: 30124358

[4] Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. Nature medicine. 2019; 25(1):24–9. https://doi.org/10.1038/s41591-018-0316-z PMID:30617335

[5] Saplacan D, Herstad J, Pajalic Z. Feedback from Digital Systems Used in Higher Education: An Inquiry into Triggered Emotions-Two Universal Design Oriented Solutions for a Better User Experience. Studies in health technology and informatics. 2018; 256:421–30. PMID: 30371503

[6] Winkler-Schwartz A, Bissonnette V, Mirchi N, Ponnudurai N, Yilmaz R, Ledwos N, et al. Artificial Intelligence in Medical Education: Best Practices Using Machine Learning to Assess Surgical Expertise in Virtual Reality Simulation. Journal of surgical education. 2019; 76(6):1681–90. https://doi.org/10.1016/j. jsurg.2019.05.015 PMID: 31202633.

[7] Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in thebrain. Psychological review. 1958; 65(6):386. https://doi.org/10.1037/h0042519 PMID: 13602029

[8] Deo RC. Machine learning in medicine. Circulation. 2015; 132(20):1920–30. https://doi.org/10.1161/CIRCULATIONAHA.115.001593 PMID: 26572668

[9] Gobet F, Lane PC, Croker S, Cheng PC, Jones G, Oliver I, et al. Chunking mechanisms in human learning.Trends in cognitive sciences. 2001; 5(6):236–43. https://doi.org/10.1016/s1364-6613(00)01662-4PMID: 11390294

[10] Ende J. Feedback in clinical medical education. Jama. 1983; 250(6):777–81. PMID: 6876333

[11] Rad, M.S., Dahlan, H.M, Iahad, N.A, Nilashi, M. and Ibrahim, O., Using a Multi-Criteria Decision-Making Approach for Assessing the Factors Affecting Social Network Sites Intention to Use. Journal of Soft Computing and Decision Support Systems 2015. 2(3): p. 20-28.

[12] Meyliana, M., Hidayanto, A.N. and Budiardjo, E.K., Evaluation of social media channel preference for student engagement improvement in universities using entropy and TOPSIS method. Journal of Industrial Engineering and Management, 2015. 8(5): p. 1676-1697.

[13] Angeli, C., Howard, S.K. Ma, J.,Yang, J. and Kirschner, P.A. , Data mining in educational technology classroom research: Can it make a contribution? . Computer and Education, 2017. 113: p. 226-242.

[14] Belachew, E.B.a.G., F.A. , Student Performance Prediction Model using Machine Learning Approach: The Case of Wolkite University.International Journal of Advanced Research in Computer Science and Software Engineering, 2017. 7(2): p. 46-50.

[15] El Mabrouk, M., Gaou, S. and Rtili, M.K. , Towards an Intelligent Hybrid Recommendation System for E-Learning Platforms Using Data Mining. International Journal of Emerging technologies in Learning,2017. 12(6): p. 52-76.

[16] Xu, J., Moon, K.H. and van der Schaar, M., A Machine Learning Approach for Tracking and Predicting Student Performance in DegreePrograms. IEEE Journal of Selected Topics in Signal Processing, 2017.11(5): p. 742-753.

[17] Khoshi, A., Gooshki, H.S. and Mahmoudi, N. , The data on the effective qualifications of teachers in medical sciences: An application of combined fuzzy AHP and fuzzy TOPSIS methods. . Data in Brief, 2018.21(2018): p. 2689-2693.

[18] Mao, Y., Lin, C. and Chi, M. , Deep Learning vs. Bayesian Knowledge Tracing: Student Models for Interventions. Journal of Educational Data Mining. , 2018. 10(2): p. 29-54.

[19] Pedro, F.; Subosa, M.; Rivas, A.; Valverde, P. Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development; UNESCO: Paris, France, 2019.

[20] Ahmad, S.F.; Rahmat, M.K.; Mubarik, M.S.; Alam, M.M.; Hyder, S.I. Artificial intelligence and its role in education. Sustainability 2021, 13, 12902. [CrossRef]

[21] Paek, S.; Kim, N. Analysis of worldwide research trends on the impact of artificial intelligence in education. Sustainability 2021 13, 7941. [CrossRef]

[22] Zheng, L.; Niu, J.; Zhong, L.; Gyasi, J.F. The effectiveness of artificial intelligence on learning achievement and learning perception: A meta-analysis. Interact. Learn. Environ. 2021, 1–15. [CrossRef]

[23] Zhang, K.; Aslan, A.B. AI technologies for education: Recent research & future directions. Compu. Edu. 2021, 2, 100025. [CrossRef]

[24] Zawacki-Richter, O.; Marín, V.I.; Bond, M.; Gouverneur, F. Systematic review of research on artificial intelligence applications in higher education: Where are the educators? Int. J. Educ. Technol. High. Ed. 2019, 16, 1–27. [CrossRef]

[25] Moltudal, S.; Høydal, K.; Krumsvik, R.J. Glimpses into real-life introduction of adaptive learning technology: A mixed methods research approach to personalised pupil learning. Design. Learn. 2020, 12, 13–28. [CrossRef]

[26] González-Calero, J.A.; Cózar, R.; Villena, R.; Merino, J.M. The development of mental rotation abilities through robotics-based instruction: An experience mediated by gender. Brit. J. Educ. Technol. 2019, 50, 3198–3213. [CrossRef]

[27] Hwang, G.-J.; Sung, H.-Y.; Chang, S.-C.; Huang, X.-C. A fuzzy expert system-based adaptive learning approach to improving students' learning performances by considering affective and cognitive factors. Compu. Edu. Art. Intel. 2020, 1, 100003. [CrossRef]

[28] Chen, L.; Chen, P.; Lin, Z. Artificial intelligence in education: A review. IEEE Access 2020, 8, 75264–75278. [CrossRef]

[29] Mou, X. Artificial intelligence: Investment trends and selected industry uses. Int. Finance. Corp. 2019, 8, 1–8.

[30] Chen, X.; Xie, H.; Hwang, G.-J. A multi-perspective study on artificial intelligence in education: Grants, conferences, journals, software tools, institutions, and researchers. Compu. Edu. 2020, 1, 100005. [CrossRef]

[31] González-Calatayud, V.; Prendes-Espinosa, P.; RoigVila, R. Artificial intelligence for student assessment: A systematic review. Appl. Sci. 2021, 11, 5467. [CrossRef]

[32] Namoun, A.; Alshanqiti, A. Predicting student performance using data mining and learning analytics techniques: A systematic literature review. Appl. Sci. 2020, 11, 237. [CrossRef]

[33] Limna, P.; Jakwatanatham, S.; Siripipattanakul, S.; Kaewpuang, P.; Sriboonruang, P. A review of artificial intelligence (AI) in education during the digital era. Adv. Know. Execu. 2022, 1, 1–9.

[34] Lameras, P.; Arnab, S. Power to the teachers: An exploratory review on artificial intelligence in education. Information 2021, 13, 14.[CrossRef]

[35] Shabbir, J.; Anwer, T. Artificial intelligence and its role in near future. J. Latex. Class. 2018, 14, 1–10. [CrossRef]

[36] Fanchamps, N.L.J.A.; Slangen, L.; Hennissen, P.; Specht, M. The influence of SRA programming on algorithmic thinking and self-efficacy using Lego robotics in two types of instruction. Int. J. Technol. Des. Educ. 2021, 31, 203–222. [CrossRef]

[37] Vanbecelaere, S.; Cornillie, F.; Sasanguie, D.; Reynvoet, B.; Depaepe, F. The effectiveness of an adaptive digital educational game for the training of early numerical abilities in terms of cognitive, noncognitive and efficiency outcomes. Brit. J. Educ. Technol. 2021, 52, 112–124. [CrossRef]

[38] Schmidt, R. A. (1991). Frequent augmented feedback can degrade learning: Evidence and interpretations. In J. Requin & G. E. Stelmach (Eds.), Tutorials in motor neuroscience (pp. 59–75). Springer.

[39] Cook DA. 2012b. Randomized controlled trials and metaanalysis in medical education: What role do they play?, Medical Teacher, In press