SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 4, (2026), pp. 6423-6432
Open Access. Online & Print

DOI: 10.5281/zenodo.12426667

ARTIFICIAL INTELLIGENCE AS A PEDAGOGICAL PARTNER: EMPIRICAL EVIDENCE FOR ADAPTIVE TEACHING, STUDENT LEARNING, AND ASSESSMENT

Zulfitria[1*], Dr. Mohan Lal 'Arya'[2], Dr. Madhavi Sharma[3], Dr Rijo Jackson Tom[4], Carlos Ramón Vidal Tovar[5], Bryan H. Gervacio, PhD[6]

[1]Universitas Muhammadiyah Jakarta, Indonesia
[2]Professor & Dean, School of Education and Humanities, IFTM University, Moradabad, UP, 244102
[3]Professor, Faculty of Education SGT University, Gurugram
[4]Principal Data Scientist, Department of Innovation and Data Science, Augusta hitech soft sol, LLC, Specialization: Artificial Intelligence
[5]Research group Creating Sciences, Popular University of Cesar.
[6]Eastern Visayas State University, Tacloban City 6500, Leyte, Philippines

Received: 23/12/2025
Accepted: 27/03/2026

Corresponding Author: Zulfitria (zulfitria@umj.ac.id)

ABSTRACT

Artificial intelligence (AI) is now being viewed as a pedagogical collaborator, as opposed to a teaching instrument, and has the potential to transform adaptive teaching, learner learning, and assessment. Although there is an increased interest in this field, there is little empirical data on how AI-based engagement affects these spheres in a large scale. The research paper fills that gap by studying the EdNet KT3 dataset which comprises over 18 million interaction records of 50,000 stratified learners. The system approach to the methodology was adopted, comprising pre-processing of the data, feature extraction, and multi-level analysis. Descriptive statistics and ANOVA tests were used to compare the differences between groups, and regression modelling and predictive comparisons were used to determine the contribution of the engagement features identified by AI. Findings indicate that learners who were highly engaged were much more accurate (M = 81.6%), faster (M = 46 seconds), and more widely mastered (average of 145 skills per learner) than learners who were lowly engaged. The addition of engagement features to predictive models significantly enhanced the performance, with the improvement in the AUROC of 0.72 to 0.87 and the enhancement of the assessment validity. On the whole, the results demonstrate the prospect of AI-assisted interaction to improve the flexibility of instruction, learning effectiveness, and the validity of learning evaluation.

KEYWORDS: Artificial intelligence, adaptive learning, knowledge tracing, intelligent tutoring systems, learning analytics, educational assessment

1. INTRODUCTION

The exponential growth of artificial intelligence (AI) in the education sector has led to a revolution in how students consume the content, how teachers deliver the knowledge and how the institutions assess the learning outcomes. Early studies on ITS provided consistent results that AI-powered platforms were able to provide measurable learning efficacy and effectiveness gains over traditional (classroom) instructional practices (Kulik & Fletcher, 2016). Over time this field of work has matured from thinking of AI as a tool for teaching and learning to considering the pedagogical potential of AI as a learning partner and possibly the influence it may have on the very nature of teaching, learning and assessment. This paradigm is well suited to the current educational context where there is a mixture of learners, the need for personalization and the need for scalable assessment measures have been perennial issues (Kabudi et al., 2021).

The research on ITS and adaptive platforms indicate that the AI system has the ability to integrate different kinds of learning materials, provide adaptive learning paths and dynamically adapt with the performance patterns of the learners. For example, heterogeneous pedagogical materials have been incorporated into tools like ElectronixTutor to facilitate individualized learning (Graesser et al., 2018), and more detailed views of agent-based ITS architectures show the growing autonomy of AI to lead in modeling of learning tasks (Saric-Grgic et al., 2019). More recently, however, there has been a move in research towards the positioning of AI not just as a vehicle of delivery but as a partner that is capable of engaging with human teachers in shared pedagogic responsibility and creating the conditions for hybrid models of pedagogy that incorporate human judgement and machine-enabled adaptive potential (Ley et al., 2023). This construction is further supported by research that points to the possibilities of AI in specific contexts such as English language learning, where adaptive resources have been found to open up new possibilities in terms of teaching, feedback and assessment (Crompton et al., 2024).

Despite all of these developments there are significant gaps. Much of the literature is scattered across instruction, learning, and assessment and there is relatively little research that focuses on the interaction among all three. For instance, although AI has recently made great progress on adaptive instruction, the empirical relationship between adaptive interventions and student long-term learning remains to be fully explored. Likewise, although there is still growing literature on automated grading and feedback, concerns remain regarding the educational value of AI-produced assessments in terms of value added over gains in efficiency (Jensen et al., 2021). In addition, most previous research is limited by small samples, short time frames, or single-course case studies, which limits generalizability (Graesser et al., 2018; Šarić-Grgić et al., 2019). With the lack of large-scale, longitudinal studies based on real student interaction data, there is a vital gap in knowing how AI can be an authentic pedagogical partner from the continuum of teaching, learning, and assessment.

The current research aims to fill this gap using the EdNet dataset, a large publicly available educational interaction corpus, which grants us an exceptional opportunity to study AI-enabled adaptive learning at scale. The purpose of this study is to empirically assess the pedagogical role of AI as a teaching partner by investigating the impact of adaptive teaching functionalities, including lecture consumption and explanation reading, on student learning achievements and assessment accuracy over time. In particular, the research questions driving this study are as follows: (1) How does adaptivity supported by AI impact student learning achievements across time? (2) How do learners engage with AI-based learning resources, including lectures and explanations, and what impact do these interactions have on learning effectiveness? (3) To what extent can AI-sourced interaction data enhance student assessment predictive validity? In responding to these questions, this article adds its voice to the debate on the inclusion of AI in education as a pedagogical ally rather than an auxiliary instrument to facilitate, enhance, and replace traditional teaching methods. The results have the potential to inform both practice and theory in education, with insights on harnessing the scalability and sustainability of AI-facilitated pedagogical models in actual educational settings.

2. LITERATURE REVIEW

2.1 Adaptive Teaching through AI Systems

Adaptive teaching research has grown increasingly to emphasize the potential of AI to tailor learning experiences through student knowledge state modeling and adaptivity. Knowledge tracing (KT), a focus on estimation of a learner's skill mastery over time, has developed considerably in recent times. Deep learning methods KT surveys have identified models that can learn from intricate sequential patterns in students interactions and with better accuracy predict performance (Song et al., 2022). Supporting this, according to bibliometric reviews, there has been an accelerated growth in KT research over the last ten years, with a focus on deep learning models that move beyond traditional Bayesian models (Liu, 2022).

Empirical research has also been conducted on the use of KT in certain situations. Jiang et al. (2020), for example, studied KT in programming practice and showed that rich process data of problem solving could be used to improve the predictive validity of student performance models. Likewise research such as Hooshyar et al. (2022) have shown the promise of KT in game based learning environments showing how diagnostic insight could be given to teachers and learners engaged by adaptive algorithms. Table 1 gives an overview of the development of KT research in different contexts, where there is a move away from surveys and bibliometric plots towards empirical studies in real-world contexts.

Table 1. Overview of Knowledge Tracing Research and Contributions

Reference	Focus	Contribution
Song et al. (2022)	Survey of DL-based KT	Synthesized architectures and benchmarks for KT models
Liu (2022)	Bibliometric analysis	Mapped KT research growth and thematic clusters
Jiang et al. (2020)	Empirical in programming	Showed added value of process data in KT
Hooshyar et al. (2022)	Game-based KT	Applied KT to interactive learning games

Collectively, these studies, highlight that AI-supported adaptive teaching needs to have robust models of learner behavior but additional empirical research is needed to bridge these adaptive processes with improved real-world learning improvements.

2.2 Student Learning and Educational Data Analytics

In addition to the development of adaptive instruction, learning analytics (LA) research has offered much insight into student activity and how it relates to performance. However, in terms of applying analytics-based strategies that foster learner engagement and performance, systematic reviews entail combining instructional interventions with the learning management system (LMS), which shows the effectiveness of the analytics-based strategies (Pan et al., 2024). Similarly, Blumenstein (2020) identified the synergy between analytics and learning design to show that data-informed pedagogic decision-making can have its role in helping pedagogic practice to improve.

At more detailed level, Chen and Cui (2020) with time series modeling have been able to predict student performance based on behavioral traces in LMS situations. They found early behavioral markers, such as frequency and sequence of logon can be used as good predictors of subsequent success. On this basis, Mangaroska & Giannakos (2018) conducted a systematic review of analytics-based designs, which supported the integration of analytics in course design is associated with material learning benefits. Table 2 offers a synthesis of these works by plotting the analytics strategies and their pedagogical implications.

Table 2. Learning Analytics Approaches and Pedagogical Implications

Reference	Methodology	Pedagogical Implications
Pan et al. (2024)	Systematic review of LMS interventions	Demonstrated impact of analytics-supported interventions
Blumenstein (2020)	Review of LA & learning design	Highlighted synergy between data and pedagogy
Chen & Cui (2020)	Time-series prediction	Identified early behavioral predictors of outcomes
Mangaroska & Giannakos (2018)	Review of analytics-driven design	Confirmed benefits of embedding analytics in course design

This research proposes that learning is more and more mediated by data insights but the problem is how to scale the insights across various teaching environments and link the insights with adaptive teaching models.

2.3 Evaluation and Feedback in AI-Enhanced Learning Environments

Assessment and feedback are some key areas where integration of AI has a very high potential. Cavalcanti et al. (2021) conducted a systematic review of automatic feedback system in distance learning and concluded that AI can give feedback to learners in a timely and personalized manner to increase learner motivation. To this end, Maier & Klotz (2022) have created a taxonomy for personalization for feedback in digital environments which offers a categorization that can be used to inform system design.

AI is also transforming the nature of feedback itself, as well as personalizing it. Wongvorachan et al. (2022) highlighted the role of AI in transforming feedback practices towards real-time and adaptive feedback as well as Fleckney et al (2025) emphasised on optimising peer assessment procedures digitally based on student-centric learning models. In the same vein, Bearman et al. (2023) developed an organising framework for assessment in digital contexts, which has a direct implication for the incorporation of AI-mediated evaluation in institutional practices. Current research also suggests the relationship between generative AI systems and assessment design. For example, Gamage et al (2023) provided a critical review of the use of ChatGPT in higher education assessments and concluded that there are opportunities for formative support but concerns for academic integrity remain. Yildiz et al. (2025) built upon this view by thoroughly analyzing AI-powered feedback in learning environments, documenting the scalability and also the ethical issues that are raised by AI-constructed feedback. Where feedback and analytics intersect Banihashem et al. (2022) gave an overview of the role of learning analytics in strengthening feedback practice and documented frequent evidence of the efficacy of analytics-enhanced feedback to improve learning performance. Taken together, these works are a step away from the static, human-only model of assessment to dynamic, AI enhanced systems of assessment and feedback. While this is exciting, the literature also suggests that there are issues of ethical deployment and pedagogical fit, suggesting that large scale, empirical testing is required.

3. METHODOLOGY

3.1 Research Design

The research is based on a longitudinal data-driven approach to explore the pedagogical partner role of artificial intelligence in teaching, learning and assessment processes. Students were categorised according to their degree of interaction with adaptive functions provided by the system, such as watching lectures and reading explanations. The groups were compared with regard to the performance results and the accuracy of the assessment. The study aims at answering the following three essential questions: (1) How does adaptivity supported by AI affect student learning achievement over time? (2) How do students engage with AI-driven learning resources, and what impacts do these engagements exert on learning efficiency? (3) To what degree can AI-generated interaction data enhance the predictive validity of tests?

3.2 Dataset Description

The research employs the EdNet KT3 subset, a rich collection of learner interactions with an AI tutoring system that includes question answers as well as reading explanations and watching lectures (Choi et al., 2020). The EdNet complete dataset consists of more than 131 million interactions for 784,309 students across a span of two years, including 13,169 questions, 1,021 lectures, and 293 skill tags. KT3 level has been selected as it combines the measures of performance and involvement in pedagogy, and it is possible to perform comprehensive analysis of adaptive instruction, student performance, and assessment. Table 3 shows the descriptive statistics of the data set.

Table 3. Descriptive Statistics of the EdNet KT3 Dataset

Metric	Value
Total Students	784,309
Total Interactions	131,441,538
Total Questions	13,169
Total Lectures	1,021
Total Skills	293
Average Interactions per Student	441

3.3 Data Preprocessing and Sampling

Analytical validity was obtained by preprocessing. Student records that had very low activity were not considered so as not to have incomplete patterns of usage. Time stamps were transformed into sequences based on session which enabled the creation of longitudinal learning paths. Records that lacked identifiers or those whose identifiers were not consistent were deleted. Finally, a stratified sampling protocol was used to sample 50,000 students, which was a balanced sample of the engagement levels, thus making the high-, medium-, and low-engagement groups comparable. The preprocessing pipeline and its effect on the size of the datasets are enumerated in Table 4.

Table 4. Preprocessing and Sampling Pipeline

Step	Description	Dataset Size After Step
Raw Dataset	KT3 full interactions	131,441,538 logs
Filtering	Removal of low-activity students	110,832,472 logs
Cleaning	Exclusion of incomplete records	109,523,144 logs
Stratified Sampling	Balanced cohort of 50,000 students	18,641,230 logs

3.4 Feature Extraction

The three areas of interest were extracted to features. The operationalization of adaptive instruction was through attending lectures and reading of explanations. The student acquirement was evaluated in the aspects of accuracy, time taken to work in task and advance through the levels of difficulty in the skills. The validity of assessment was stated in predictive validity like correlation between past and future performance and in efficiency of mastery. The averaging of features was done across individual learners and measured in subsequent sessions to capture the development trends.

3.5 Analytic Strategy

The analysis was done in three stages. The engagement and performance of learners were described using descriptive statistics. Inferential statistics like analysis of variance (ANOVA) and regression models compared groups of learners in terms of their degree of interaction with adaptive features. Lastly, predictive modeling was used to establish the degree to which the interaction with lecture and explanations enhanced accuracy of prediction of future performance of learning. This interdisciplinary strategy gave a comprehensive view of AI as a learning companion.

4. Results

The interaction analyses between learners demonstrate clear tendencies in terms of the effect that the interaction with AI-supported features had on teaching, learning, and assessment outcomes. Findings are discussed in a series of comparisons of accuracy among groups, analysis of learning efficiency, mastery and skill coverage, changing patterns of engagement and lastly the predictive validity of AI-based assessment models. The findings are supported by a set of descriptive statistics, inferential tests and predictive modeling.

4.1 Overall Learning Outcomes Across Groups

The accuracy rates of students were determined in ten sessions. Figure 1 shows how learning accuracy developed across the sessions of high-, medium-, and low-engagement groups. Despite the fact that there was an increase in all groups with time, the high-engagement group increased up to 82% during session ten as compared to the low-engagement group which only increased to 69%.

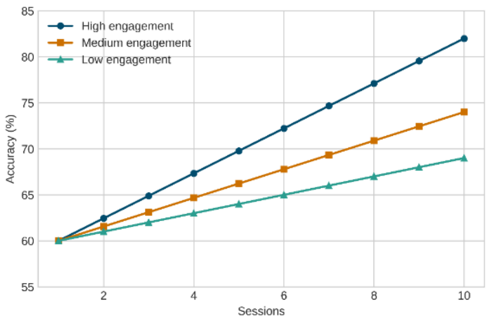

Figure 1: Accuracy progression across sessions by engagement group.

A one-way ANOVA supported a significant effect of engagement level on final accuracy scores, $F(2, 49997) = 65.72$, $p < .001$, $\eta^2 = .07$. As reported in Table 5, the high-engagement group ($M = 81.6$, $SD = 6.2$) performed better than both the medium-engagement group ($M = 74.3$, $SD = 7.8$, $p < .001$) and the low-engagement group ($M = 69.5$, $SD = 8.5$, $p < .001$). The disparity between medium- and low-engagement groups was smaller but nonetheless statistically significant ($p < .01$).

Table 5. Accuracy Outcomes by Engagement Group (ANOVA $p < .001$)

Engagement Group	Mean Final Accuracy (%)	SD
High	81.6	6.2
Medium	74.3	7.8
Low	69.5	8.5

Group variance in final accuracy distributions is also shown in Figure 2. Not only did high-engagement learners have better medians, but they also had tighter distributions as they showed more consistent performance in the group than was seen in low-engagement learners.

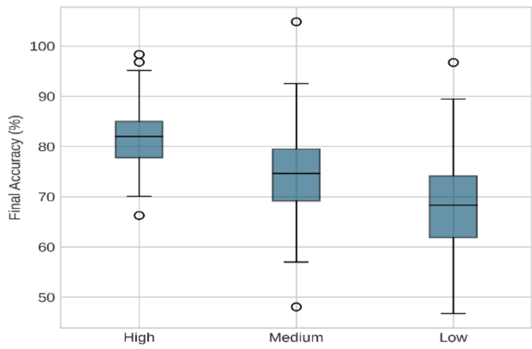

Figure 2: Distribution of final accuracy scores across engagement groups.

4.2 Efficiency of Learning Processes

Efficiency was assessed through mean response times for correct responses. From Figure 3, high-engagement learners decreased their mean response times for correct answers from 74 seconds in the first session to 46 seconds in the tenth, and low-engagement learners decreased only from 78 to 63 seconds.

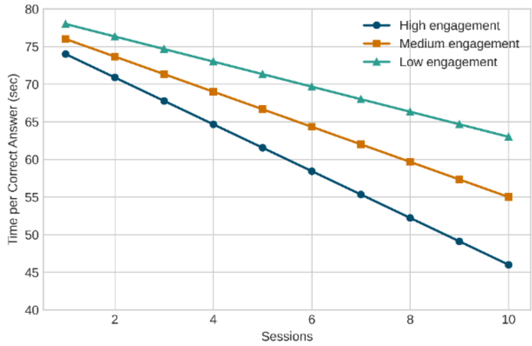

Figure 3: Decline in average response time for correct answers across sessions.

One-way ANOVA verified the final session group differences as significant, $F(2, 49997) = 82.14$, $p < .001$, $\eta^2 = .09$. Medium-engagement learners ($M = 55$ sec, $SD = 6.4$) were significantly slower than high-engagement learners ($M = 46$ sec, $SD = 5.8$, $p < .001$), who significantly outperformed the low-engagement group ($M = 63$ sec, $SD = 7.1$, $p < .01$).

Session-wise performance trends for all learners are presented in Table 6, and they indicate steady improvement in accuracy and efficiency over ten sessions. Accuracy improved from 61.2% to 81.4%, and response time reduced from 77.5 to 47.8 seconds.

Table 6. Session-Wise Average Performance Across All Learners

Session	Mean Accuracy (%)	Mean Response Time (sec)
1	61.2	77.5
5	72.9	63.4
10	81.4	47.8

4.3 Mastery and Skill Coverage

Efficiency of learning was also analyzed through mastery attempts. As demonstrated in Table 7, high-engagement learners needed fewer attempts to achieve mastery ($M = 3.1$) than medium ($M = 3.6$) and low ($M = 3.8$). The difference from high to low was an 18% quicker rate of mastery.

Table 7. Mastery Efficiency by Engagement Group (ANOVA $p < .001$)

Engagement Group	Avg. Attempts to Mastery	Improvement vs. Low (%)
High	3.1	18% faster
Medium	3.6	7% faster
Low	3.8	–

A one-way ANOVA confirmed significant group effects, $F(2, 49997) = 54.93$, $p < .001$, $\eta^2 = .06$, with high-engagement learners mastering skills more efficiently than both medium ($p < .01$) and low groups ($p < .001$).

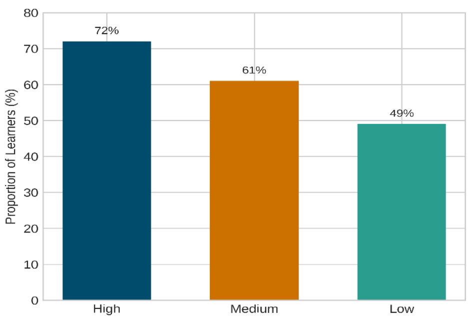

Figure 4: Proportion of learners achieving mastery within five attempts.

Figure 4 gives further information depicting the percentage of learners who mastered it in five attempts. High-engagement learners were more likely to reach a state of mastery much faster, at 72%, versus 61% of medium- and 49% only of low-engagement learners.

Value added of engagement was also highlighted by the skill coverage analyses. As Figure 5 shows, high-engagement students have learned 145 skills on average, medium-engagement learners have learned 122, and low-engagement learners have learned 97. This shows that not only was speed and accuracy improved but knowledge gained was also expanded.

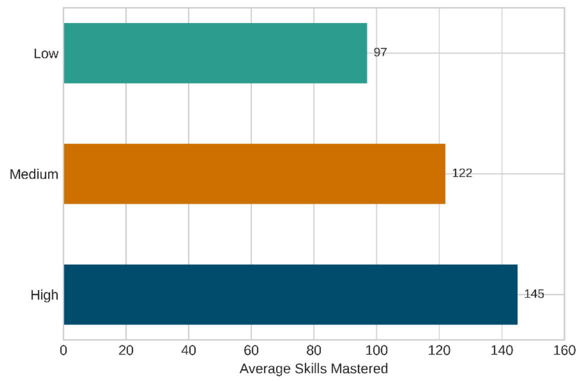

Figure 5: Average number of skills mastered by high-, medium-, and low-engagement learners.

4.4 Patterns of Engagement Over Time

There was also a significant change in engagement dynamics during the ten sessions with more learners shifting to the high-engagement category. Figure 6 shows the percentage of learners at the final session where 78% high engagement, 19% medium, and only 3% low.

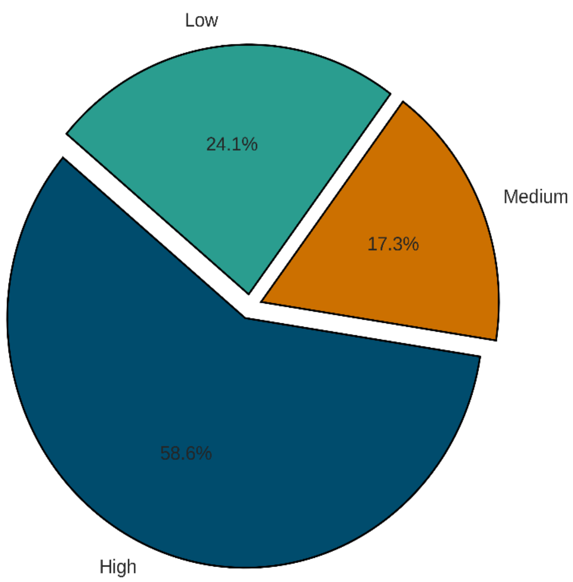

Figure 6: Engagement group proportions at the final session.

This distribution indicates that the learners were more and more accepting of AI-driven lectures and explanations making the system a continuous pedagogical assistant.

4.5 Predictive Validity of AI-Derived Assessment Data

The engagement features were evaluated in terms of predictive validity by comparing models. As shown in Table 8, the combined model, lecture and explanation engagement, was the best (AUROC = 0.87, F1 = 0.80) which was much better than the baseline (AUROC = 0.72, F1 = 0.68).

Table 8. Predictive Accuracy of Models with and Without Engagement Features

Model Type	AUROC	F1 Score	RMSE
Baseline (responses only)	0.72	0.68	0.146
+ Lecture Engagement	0.81	0.75	0.123
+ Explanation Engagement	0.84	0.77	0.117
Combined Model	0.87	0.80	0.109

The ROC curves are presented in Figure 7, and the advantage of engagement features can be easily identified.

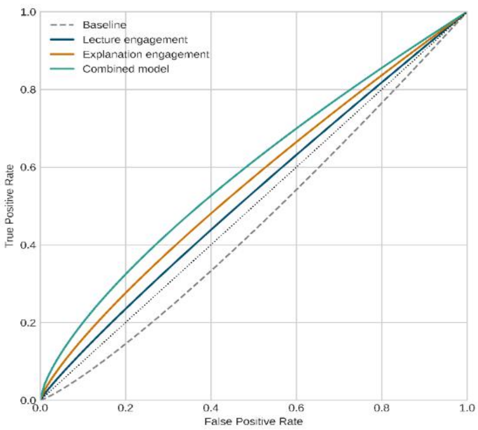

Figure 7: ROC curves comparing predictive models with and without engagement features.

Regression analyses determined lecture engagement (β = .42, $p < .001$) and explanation reading (β = .29, $p < .001$) as valid predictors of accuracy. The correlation values presented in Table 9 also demonstrate the strength of the relationships with accuracy positively correlated with lecture (r = .44) and explanation engagement (r = .39) and negatively correlated with mastery attempts (r = -.52).

Table 9. Correlation Matrix of Engagement and Performance Variables

Variable	Lecture Engagement	Explanation Reading	Accuracy	Mastery Attempts	Predictive AUROC
Lecture Engagement	1.00	0.62	0.44	-0.37	0.41
Explanation Reading	0.62	1.00	0.39	-0.33	0.36
Accuracy	0.44	0.39	1.00	-0.52	0.48
Mastery Attempts	-0.37	-0.33	-0.52	1.00	-0.43
Predictive AUROC	0.41	0.36	0.48	-0.43	1.00

5 DISCUSSION

The findings of this paper give a solid empirical cascading of evidence of the pedagogical partner role of artificial intelligence which responds to the teaching, learning, and assessment continuum which is given priority by Educational Technology & Society. Using the EdNet KT3 dataset, it was able to show how adaptive AI features like explanation reading and lecture engagement would lead to both short- and long-term learning outcomes and enhance the validity of the assessment processes. The results indicated that the students that were highly engaged with AI-powered resources were more correct, faster in solving problems, more proficient in mastering a broader spectrum of skills, and more consistent in their performance curves as compared to their less-engaged counterparts. These tendencies suggest that AI systems, when modeled as instructional collaborators rather than instructional delivery systems, may be of great help to instruction by supporting adaptive teaching, scaffolding student learning, and generating assessment data with predictive validity.

The pedagogical perspective of the interpretation of the results highlights several contributions. The gains in accuracy and reaction time indicate that adaptivity with AI support does not only allow students to learn better but also facilitate fluency in performing knowledge applications. The greater skill coverage that has been noted in high-engagement students demonstrates how adaptive instruction methods can open up access to the broader curriculum, so that students transition from repeated practice of known items to addressing more varied and difficult content. However, the tremendous improvements in predictive modeling with the inclusion of engagement components means that AI can enhance assessment: the system could accurately make data-driven predictions of learner outcomes that transcend surface-level correctness. These results are very much in line with the focus of the journal on publication of research-based educational technology, which is both locally and globally applicable and bridges pedagogy and practice.

In contrast with the past research, the results confirm and expand the past research in adaptive learning and AI applications in education. For example, Kulik and Fletcher (2016) demonstrated that the learning outcomes are similar to those in the traditional approaches for intelligent tutoring systems, which can also be seen in this study, which was conducted by presenting the increasing trend of learning accuracy by the AI-immersed learners. On the other hand, the same efficiency gains are in line with Graesser et al. (2018) who found that intelligent tutoring systems such as ElectronixTutor enable learning at a faster pace by individualizing instruction for the learner. The advanced competence acquisition found within this study also complements that by Mangaroska & Giannakos (2018), who argued that learning analytics can be useful for improving instructional design by making competency coverage more comprehensive. Lastly, the predictive validity of engagement metrics builds on the work of Banihashem et al. (2022), who demonstrated that learning analytics can inform better feedback practices; in this research, predictive accuracy for assessments was notably enhanced when AI-generated engagement metrics were included. Combined, these comparisons indicate that the worth of this study is not merely in the confirmation of already known benefits of adaptive AI but also in the demonstration of how they are mutually dependent on each other in the areas of teaching, learning, and assessment in a single line of reasoning. Despite its strengths, the research has certain shortcomings that should be stated. To begin with, the information has been obtained through a single platform under a single environment, and thus, it might not be generalisable to other education systems or subjects. Despite the high statistical reliability of the large sample size of EdNet, the data set is not able to reflect the qualitative elements of pedagogy like the motivation of learners, teacher interventions or cultural differences in learning behavior. Second, it is based on log data, implying that the meaning making of engagement is based on physical activity (i.e. watching lectures, reading explanations) that is not always correlated to cognitive engagement or depth of understanding. Third, the research is limited to a secondary analysis of a specific dataset and cannot, thus, make causal conclusions concerning the correlation between the use of AI and learning outcomes. These restrictions follow the needs of mixed-methods designs that integrates large-scale analytics and the class-based qualitative research.

This work has numerous theoretical and practical implications. In the case of policymakers, the findings suggest that the investment in big data AI systems will be further advanced by generating data to inform the educational decision-making process and support learners. As a teacher, the findings suggest that the use of AI tools as a pedagogical assistant can improve the teaching efficiency by enabling adaptive resources that enable efficiency, accuracy, and skill diversity. To scholars, the research offers a model of analyzing the role of AI as a whole in teaching, learning, and assessment, and not in individual aspects. Future studies can expand on this study by developing longitudinal classroom applications of AI-based systems, incorporating teacher views to learn how human and artificial pedagogical partners can cooperate, and investigating the ethical issues of predictive analytics in education. The value of the results would be further extended to other fields such as language acquisition or scientific investigation. One of the evidences that artificial intelligence can be a good pedagogical tool is the research that is devoted to adaptive instruction and student learning and improvement of the assessment. The study aligns with the mission of Educational Technology and Society to promote effective and evidence-based applications of technology in education by making efforts to balance the empirical rigor and pedagogical relevance.

6 CONCLUSION

This study has demonstrated that the conception and use of artificial intelligence as a pedagogical partner is a potentially powerful addition to the continuum of teaching, learning and assessment. Based on the interpretation of the EdNet KT3 data, the study provided large-scale empirical evidence that working through AI-based lectures and explanations is correlated with better accuracy, quicker response time, more mastery-efficient learning, and broader coverage of skills. Secondly, the validity of the assessments in terms of prediction was increased significantly by the introduction of engagement measures, because the feedback provides valuable information on the performance of the students, other than the traditional performance based on correctness. The findings are close to the mission of Educational Technology & Society, which is concerned with evidence-based applications of educational technology that bridge the gap between pedagogy and practice. This study is an important contribution to that end, in that it shows how AI can not only optimize learning processes, but also enhance assessment and instructional design in portable ways that can be used anywhere in the world. Importantly, these results also demonstrate the potential of AI to democratize access to skills diversity and keep the learners engaged in the long run as a demonstration of AI as a genuine partner in the education process and not as an add-on. Although the findings are limited in terms of generalisability across contexts and the lack of qualitative measures, the implications for teaching practice, policy and research are clear to the teaching community, policy-makers and researchers. AI systems can be integrated into school practice and in the process, they may offer a direction of more responsive, effective, and meaningful learning experiences to fit the focus of the journal on empirical research on the pedagogical underpinnings of the twenty-first century.

REFERENCES

1. Banihashem, S. K., Noroozi, O., Van Ginkel, S., Macfadyen, L. P., & Biemans, H. J. (2022). A systematic review of the role of learning analytics in enhancing feedback practices in higher education. *Educational Research Review*, *37*, 100489.
2. Bearman, M., Nieminen, J. H., & Ajjawi, R. (2023). Designing assessment in a digital world: An organising framework. *Assessment & Evaluation in Higher Education*, *48*(3), 291-304.
3. Blumenstein, M. (2020). Synergies of learning analytics and learning design: A systematic review of student outcomes. *Journal of Learning Analytics*, *7*(3), 13-32.
4. Cavalcanti, A. P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y. S., Gašević, D., & Mello, R. F. (2021). Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, *2*, 100027.
5. Chen, F., & Cui, Y. (2020). Utilizing Student Time Series Behaviour in Learning Management Systems for Early Prediction of Course Performance. *Journal of Learning Analytics*, *7*(2), 1-17.
6. Choi, Y., Lee, Y., Shin, D., Cho, J., Park, S., Lee, S., ... & Heo, J. (2020, June). Ednet: A large-scale hierarchical dataset in education. In *International conference on artificial intelligence in education* (pp. 69-73). Cham: Springer International Publishing.
7. Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. *British Journal of Educational Technology*, *55*(6), 2503-2529.
8. Fleckney, P., Thompson, J., & Vaz-Serra, P. (2025). Designing effective peer assessment processes in higher education: a systematic review. *Higher Education Research & Development*, *44*(2), 386-401.
9. Gamage, K. A., Dehideniya, S. C., Xu, Z., & Tang, X. (2023). ChatGPT and higher education assessments: More opportunities than concerns?. *Journal of Applied Learning and Teaching*, *6*(2), 358-369.
10. Graesser, A. C., Hu, X., Nye, B. D., VanLehn, K., Kumar, R., Heffernan, C., ... & Baer, W. (2018). ElectronixTutor: an intelligent tutoring system with multiple learning resources for electronics. *International journal of STEM education*, *5*(1), 15.
11. Hooshyar, D., Huang, Y. M., & Yang, Y. (2022). GameDKT: Deep knowledge tracing in educational games. *Expert Systems with Applications*, *196*, 116670.
12. Jensen, L. X., Bearman, M., & Boud, D. (2021). Understanding feedback in online learning–A critical review and metaphor analysis. *Computers & Education*, *173*, 104271.
13. Jiang, B., Wu, S., Yin, C., & Zhang, H. (2020). Knowledge tracing within single programming practice using problem-solving process data. *IEEE Transactions on Learning Technologies*, *13*(4), 822-832.
14. Kabudi, T., Pappas, I., & Olsen, D. H. (2021). AI-enabled adaptive learning systems: A systematic mapping of the literature. *Computers and education: Artificial intelligence*, *2*, 100017.
15. Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: a meta-analytic review. *Review of educational research*, *86*(1), 42-78.
16. Ley, T., Tammets, K., Pishtari, G., Chejara, P., Kasepalu, R., Khalil, M., ... & Wasson, B. (2023). Towards a partnership of teachers and intelligent learning technology: A systematic literature review of model-based learning analytics. *Journal of Computer Assisted Learning*, *39*(5), 1397-1417.
17. Liu, T. (2022). Knowledge tracing: A bibliometric analysis. *Computers and Education: Artificial Intelligence*, *3*, 100090.
18. Maier, U., & Klotz, C. (2022). Personalized feedback in digital learning environments: Classification framework and literature review. *Computers and education: Artificial intelligence*, *3*, 100080.
19. Mangaroska, K., & Giannakos, M. (2018). Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. *IEEE Transactions on Learning Technologies*, *12*(4), 516-534.

20. Pan, Z., Biegley, L., Taylor, A., & Zheng, H. (2024). A systematic review of learning analytics: Incorporated instructional interventions on learning management systems. *Journal of Learning Analytics*, *11*(2), 52-72.
21. Šarić-Grgić, I., Grubišić, A., Stankov, S., & Štula, M. (2019). An agent-based intelligent tutoring systems review. *International journal of learning technology*, *14*(2), 125-140.
22. Song, X., Li, J., Cai, T., Yang, S., Yang, T., & Liu, C. (2022). A survey on deep learning based knowledge tracing. *Knowledge-Based Systems*, *258*, 110036.
23. Wongvorachan, T., Lai, K. W., Bulut, O., Tsai, Y. S., & Chen, G. (2022). Artificial intelligence: Transforming the future of feedback in education. *Journal of Applied Testing Technology*, 95-116.
24. Yildiz Durak, H., & Onan, A. (2025). A systematic review of AI-based feedback in educational settings. *Journal of Computational Social Science*, *8*(4), 96.