

RESEARCH

Received: December 18, 2025

Revised: January 21, 2026

Accepted: February 7, 2026

Effectiveness of AI-Based Deep Learning Modules for Vocational Students' Mathematical Logic Understanding

Ema Yuliantri^{1*}, Sudarwan Danim², Nurul Astuti Yensy B.³

***Correspondence:**

Ema Yuliantri
ema.yuliantri80@gmail.com

¹Doctoral Study Program in Education, Faculty of Teacher Training and Education (FKIP), University of Bengkulu, Indonesia

²Doctoral Study Program in Education, Faculty of Teacher Training and Education (FKIP), University of Bengkulu, Indonesia

³Doctoral Study Program in Education, Faculty of Teacher Training and Education (FKIP), University of Bengkulu, Indonesia

Abstract

Artificial intelligence (AI) can support adaptive mathematics learning through personalized sequencing, immediate feedback, and analysis of student errors. This study examined the effectiveness of an AI-based deep-learning module for improving vocational high school students' understanding of mathematical logic. A quasi-experimental pretest-posttest control-group design involved 60 Grade 10 students from one public vocational high school in Indonesia. The experimental class learned with the AI-supported module for six meetings, while the control class received regular instruction. Mathematical logic understanding was measured using a 20-item test combining essay and reasoned multiple-choice items; the instrument was validated and showed acceptable reliability ($\alpha > .70$). Descriptive statistics, normalized gain, and independent-samples t-tests were used. The experimental class improved from $M = 56.42$ to $M = 82.15$ ($N\text{-gain} = .62$), whereas the control class improved from $M = 55.87$ to $M = 71.03$ ($N\text{-gain} = .38$). Both gains fall within the conventional medium category, but the experimental group showed a significantly larger improvement ($p < .05$). The findings indicate that AI-supported adaptive modules can strengthen conceptual understanding of mathematical logic when integrated with structured instruction and timely feedback.

Keywords: adaptive learning; artificial intelligence; deep learning module; mathematical logic; vocational high school

Introduction

The rapid development of artificial intelligence (AI) has reshaped educational technology by enabling systems to provide personalized guidance, predictive support, automated feedback, and adaptive sequencing of learning activities. Contemporary AI in education (AIEd) is therefore not limited to automating administrative tasks; it increasingly functions as an instructional layer that can respond to learners' performance and support teachers in making evidence-informed decisions (Hwang et al., 2020; Ouyang & Jiao, 2021; Zhai et al., 2021). Systematic reviews show that AIEd research has expanded rapidly across intelligent tutoring systems, adaptive learning, learning analytics, automated assessment, and generative AI, although the pedagogical quality of implementation remains as important as the sophistication of the technology itself (Chen et al., 2020; Chiu et al., 2023; Zawacki-Richter et al., 2019). In mathematics education, this development is particularly relevant because abstract concepts often require repeated practice, immediate

diagnosis of misconceptions, and feedback that is sensitive to students' prior knowledge and learning pace.

Evidence from mathematics and technology-enhanced learning suggests that adaptive systems can improve achievement when they are integrated into coherent instruction rather than used as stand-alone digital replacements. Meta-analyses of intelligent tutoring systems have reported positive learning effects across educational levels and subject areas, including mathematics (Kulik & Fletcher, 2016; Ma et al., 2014; Steenbergen-Hu & Cooper, 2013). A large meta-analysis of digital tools in secondary mathematics and science also found that technology-supported instruction produced meaningful learning gains, with intelligent tutoring systems and simulations among the more effective categories (Hillmayr et al., 2020). More recent reviews in mathematics education similarly indicate that AI-based tutoring, adaptive feedback, and personalized task sequencing can support conceptual understanding and problem solving, although effect sizes vary according to implementation quality, learner characteristics, and teacher readiness (Hwang, 2022; Molemane et al., 2026; Son, 2024). This pattern suggests that AI should be understood as a pedagogical support mechanism whose value depends on how well adaptivity, feedback, and learning tasks are aligned with intended mathematical competencies.

Within vocational high schools, the instructional challenge is especially important because mathematics is expected to support occupational reasoning, problem solving, programming, data interpretation, and decision making rather than operate only as an academic subject. Vocational education is oriented toward the development of knowledge and capabilities that can be transferred to authentic work practices (Billett, 2011). Mathematical logic contributes to this orientation because propositions, truth values, implication, equivalence, and structured reasoning provide foundations for algorithmic thinking and systematic problem solving. The broader mathematics education literature also emphasizes that conceptual understanding and mathematical reasoning require students to connect representations, justify conclusions, and monitor the validity of solution processes rather than merely reproduce procedures (Kilpatrick et al., 2001; Lithner, 2008; Schoenfeld, 1985). For vocational learners, these capacities are increasingly intertwined with digital and computational competencies.

Mathematical logic is also a suitable domain for studying adaptive learning because student errors are often diagnostic rather than random. Incorrect responses can reveal whether a learner misunderstands negation, confuses necessary and sufficient conditions, misreads symbolic notation, or applies an invalid inference rule. This makes it possible for a digital system to connect observable response patterns with targeted follow-up tasks. Research on mathematical reasoning emphasizes that productive learning requires learners to move beyond imitation toward explanation, justification, and strategic choice (Lithner, 2008; Schoenfeld, 1985). From this perspective, the value of an AI-supported module lies not in accelerating answer production but in creating repeated opportunities for students to inspect why an argument is valid or invalid. The same principle is reflected in work on intelligent tutoring systems, where the strongest educational designs model learner knowledge, provide hints or feedback at appropriate moments, and preserve opportunities for students to perform the reasoning themselves (Kulik & Fletcher, 2016; VanLehn, 2011). For vocational students, such structured reasoning has practical relevance because technical procedures frequently depend on conditional decisions: if a condition is met, a process continues; if not, an alternative action is required. Strengthening formal logic can therefore support both school mathematics and the broader habits of systematic decision making required in technical work.

The term deep learning can create conceptual ambiguity because it is used in both artificial intelligence and pedagogy. In computer science, deep learning normally refers to multilayer neural-

network methods, whereas in educational discourse it may refer more generally to learning that promotes conceptual connections, reasoning, reflection, and transfer. In this study, the phrase AI-based deep learning module is used operationally to describe an AI-supported instructional module that promotes cognitively deep engagement through adaptive presentation of content, automated feedback, and analysis of student error patterns. The study does not claim to evaluate a specific neural-network architecture. This operational clarification is important because the educational effectiveness of an AI system should be judged by the quality of its instructional functions and learning outcomes, not by the technological label alone (Guan et al., 2020; Kabudi et al., 2021).

Three mechanisms provide a theoretical rationale for expecting an AI-supported module to improve mathematical logic understanding. First, adaptive learning systems can vary task sequence, support, and content in response to learner performance, thereby reducing the mismatch between a fixed curriculum pace and heterogeneous student readiness (Kabudi et al., 2021; Nurhayati et al., 2026). Second, well-designed feedback can focus attention on the task, the process used to solve it, and strategies for self-correction. Research on formative feedback consistently shows that timely, specific, and actionable information is more useful than feedback that only indicates whether an answer is right or wrong (Hattie & Timperley, 2007; Shute, 2008). Third, multimedia and adaptive instructional design can reduce unnecessary cognitive processing while directing attention toward relevant conceptual relationships, provided that the module avoids visual and informational overload (Mayer, 2021; Sweller et al., 2019).

The current evidence base also indicates a need for context-specific studies. Much AIED research has been concentrated in higher education, while secondary and vocational settings are less represented (Crompton & Burke, 2023; Zawacki-Richter et al., 2019). Indonesian vocational teachers generally report positive attitudes toward AI as a pedagogical partner, but adoption still depends on infrastructure, teacher competence, and clear instructional purposes (Yuliansah et al., 2026). Recent Indonesian reviews of adaptive and generative AI in mathematics likewise emphasize personalization and responsive feedback while noting the importance of teacher preparedness and pedagogical integration (Maryati et al., 2026; Nurhayati et al., 2026). These findings justify empirical work that examines not simply whether AI is present in a classroom, but whether a specific AI-supported learning design produces measurable improvement in a difficult mathematical topic.

Accordingly, the novelty of the present study is its focus on an AI-supported module for mathematical logic in the vocational secondary-school context. Rather than examining general attitudes toward AI or broad mathematics achievement, the study evaluates changes in a defined conceptual domain using a quasi-experimental pretest-posttest comparison. The study asks whether students who learn through an AI-supported deep-learning module show greater improvement in mathematical logic understanding than students who receive regular instruction. It also considers the instructional mechanisms that may account for the observed differences, particularly adaptive content, automatic feedback, and analysis of recurring errors. The findings are expected to contribute to the evidence base on AI-supported mathematics learning while providing practical guidance for vocational teachers and learning-material developers.

Method

This study employed a quantitative quasi-experimental pretest-posttest control-group design. A quasi-experimental design was appropriate because the research was conducted in an authentic school setting where complete individual randomization was not feasible, while a comparison group and pretest were retained to strengthen causal interpretation (Creswell & Creswell, 2018;

Shadish et al., 2002). The participants were 60 Grade 10 students enrolled at one public vocational high school in Indonesia and distributed across an experimental class and a control class. Purposive sampling was used to select classes with comparable curricular exposure and initial mathematical learning conditions. Because intact school classes were involved, the design is interpreted as a classroom-based effectiveness study rather than a fully randomized trial.

The intervention was implemented for six instructional meetings. Before treatment, both groups completed the same pretest of mathematical logic understanding. The experimental class then used the AI-supported module, whereas the control class continued with regular teacher-led instruction. The module was designed around three functions already embedded in the instructional product: adaptive presentation of learning content, automatic feedback on student responses, and analysis of recurring error patterns to inform subsequent practice. These functions are consistent with the broader architecture of AI-enabled adaptive learning systems, in which learner data are used to modify support and task progression (Heffernan et al., 2016; Kabudi et al., 2021). At the end of the intervention, both groups completed the same posttest under comparable conditions. The relatively short six-meeting period was chosen to evaluate immediate learning effects and should not be interpreted as evidence of long-term retention.

To maintain comparability across the intervention period, the same mathematical logic competency domain and assessment targets were used for both groups. The experimental treatment differed primarily in the way content and feedback were delivered through the AI-supported module, while the control condition represented the regular instructional approach already used in the school. In an effectiveness study of educational technology, this distinction is important because learning gains may be influenced by the novelty of the medium, time-on-task, teacher assistance, and alignment between practice activities and the test. The study therefore interprets the module as an integrated instructional treatment rather than attributing outcomes to a single AI algorithm. This approach is consistent with research showing that technology effects in mathematics are moderated by implementation conditions and by how well the digital tool is integrated with instruction (Hillmayr et al., 2020; Pane et al., 2014). During implementation, the teacher retained responsibility for classroom management, clarification of directions, and support when students encountered technical or conceptual difficulties. Maintaining teacher involvement is also consistent with contemporary AIED frameworks that position AI as a support for human teaching and learner agency rather than as a replacement for professional judgment (Ouyang & Jiao, 2021).

The mathematical logic understanding test consisted of 20 items combining essay questions and reasoned multiple-choice questions. The use of reasoned responses was intended to capture not only answer accuracy but also students' ability to justify logical relationships and identify the reasoning underlying a conclusion. The instrument was developed from the instructional competency indicators used in the module and reviewed through expert judgment for relevance, clarity, and alignment. Empirical item validity was examined using item-total correlations, while internal consistency was estimated using Cronbach's alpha. The reported reliability coefficient exceeded .70, which was treated as acceptable for the present classroom research; however, alpha was interpreted as evidence of internal consistency rather than as a guarantee of unidimensionality or construct validity (Azwar, 2020; Taber, 2018).

The data analysis proceeded in three stages. First, descriptive statistics were used to summarize pretest and posttest performance. Second, normalized gain was calculated to describe improvement relative to the maximum possible score. Normalized gain was interpreted using the conventional categories introduced by Hake (1998): low ($< .30$), medium (.30-.69), and high ($\geq .70$).

.70). Third, inferential comparison between groups used independent-samples t-tests after checking the assumptions of normality and homogeneity, with statistical significance set at $p < .05$ (Field, 2018). Because the manuscript dataset available for this article contains summary scores rather than the complete student-level dataset, exact confidence intervals and standardized effect sizes are not reconstructed here. The reported inferential conclusion is therefore presented together with the descriptive and normalized-gain evidence rather than being overstated.

Ethical considerations were integrated into the research procedure. School approval was obtained before data collection, students were informed about the learning activities and research use of their data, and participation was handled in accordance with informed-consent and confidentiality principles. Student identities are not reported in the manuscript. These procedures are consistent with established educational research ethics concerning voluntary participation, privacy, and responsible handling of learner data (British Educational Research Association, 2018). In AI-supported learning contexts, protecting student data and maintaining meaningful teacher oversight are also important because personalization depends on the collection and interpretation of performance information (Chiu et al., 2023; Hwang et al., 2020).

Results

The descriptive results show that the experimental and control classes began at similar mean levels of mathematical logic understanding. The experimental class had a mean pretest score of 56.42, while the control class had a mean of 55.87, a difference of only 0.55 points. After six meetings, the experimental mean increased to 82.15, whereas the control mean increased to 71.03. Thus, the absolute score gain was 25.73 points in the experimental class and 15.16 points in the control class. The pattern suggests that both groups learned during the intervention period, but the class using the AI-supported module showed a substantially larger improvement.

Table 1. Mean Pretest, Posttest, and N-Gain Scores for Mathematical Logic Understanding

Group	Pretest (M)	Posttest (M)	N-Gain	Category
Experimental Class	56.42	82.15	0.62	Medium
Control Class	55.87	71.03	0.38	Medium

Table 1 shows that the experimental class achieved a normalized gain of 0.62, compared with 0.38 in the control class. Under the conventional Hake (1998) interpretation, both values fall within the medium category, but the experimental gain is markedly closer to the upper boundary of that category and is 0.24 points higher than the control gain. The reported independent-samples t-test indicated a statistically significant between-group difference ($p < .05$). This combination of evidence supports the conclusion that the AI-supported module was associated with stronger improvement in mathematical logic understanding during the study period. The result should be interpreted as evidence of short-term classroom effectiveness rather than proof that AI alone caused the improvement, because the design used intact classes and the available summary does not provide a standardized effect size.

The improvement pattern is educationally meaningful because the module targeted processes that are central to learning mathematical logic: identifying relationships among propositions, checking the validity of implications, recognizing errors in reasoning, and receiving corrective information before misconceptions became stable. Research on intelligent tutoring and adaptive mathematics systems suggests that learning gains are more likely when feedback is connected to specific learner errors and when practice is sequenced according to demonstrated understanding (Ma et al., 2014; Roschelle et al., 2016; VanLehn, 2011). The present results are consistent with that

mechanism, although the study did not separately estimate the contribution of each module feature.

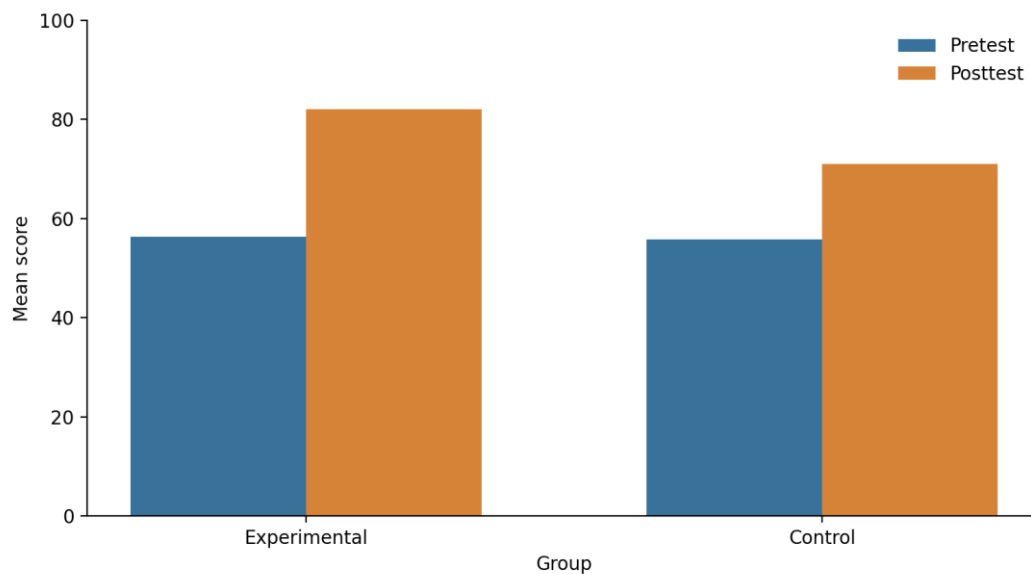


Fig. 1 Pretest and posttest mean scores by group

Discussion

The higher improvement observed in the experimental class is consistent with a substantial body of research showing that AI-supported and intelligent tutoring environments can strengthen mathematics learning when they provide individualized support rather than simply digitizing conventional exercises. Meta-analytic evidence indicates that intelligent tutoring systems generally outperform conventional large-group instruction and non-intelligent computer-based instruction, although effects differ by implementation and comparison condition (Kulik & Fletcher, 2016; Ma et al., 2014). In K-12 mathematics specifically, AI and computer-based interventions have produced positive average effects, including among lower-performing learners (Hwang, 2022; Ran et al., 2021; Steenbergen-Hu & Cooper, 2013). The present study extends this pattern to mathematical logic in a vocational high-school setting, where empirical evidence remains comparatively limited.

One plausible explanation is the quality and timing of feedback. Mathematical logic problems often expose misconceptions immediately: a student may confuse converse and inverse statements, treat implication as equivalence, or apply truth-table rules inconsistently. In a conventional classroom, these errors may not be identified until assignments are reviewed. An AI-supported module can potentially shorten this feedback cycle by detecting incorrect response patterns and returning task-specific support. The broader feedback literature shows that feedback is most effective when it helps learners understand where they are, what the goal is, and what action should be taken next (Hattie & Timperley, 2007; Shute, 2008). Intelligent tutoring research similarly emphasizes step-level or response-sensitive guidance as a major mechanism through which digital tutors approximate individualized instructional support (VanLehn, 2011).

A second explanation concerns adaptivity. Fixed instruction assumes that all learners require the same examples, pace, and level of support. Adaptive systems attempt to use evidence from learner performance to vary content, practice, or guidance. Systematic mapping studies have shown that AI-enabled adaptive learning systems commonly use learner data to support personalization, while recent mathematics-focused reviews report generally positive outcomes when adaptivity is pedagogically integrated and teachers remain actively involved (Kabudi et al., 2021; Nurhayati et al., 2026). The experimental group's larger gain is therefore compatible with the view that students

benefit when practice is better matched to their current understanding. However, adaptivity should not be equated with unrestricted automation. Ouyang and Jiao (2021) argue that the development of AIED is moving from AI-directed learning toward models in which AI supports or empowers learners while preserving learner agency.

The result can also be interpreted through principles of cognitive and multimedia learning. Mathematical logic requires learners to coordinate symbolic forms, verbal statements, truth conditions, and inferential rules. Poorly designed digital materials can increase extraneous cognitive load, whereas structured representations and well-timed cues can help students focus limited working-memory resources on the relationships that matter (Mayer, 2021; Sweller et al., 2019). Adaptive sequencing may also prevent students from moving too quickly to complex items before prerequisite concepts are sufficiently stable. This does not mean that AI automatically reduces cognitive load; rather, instructional design determines whether the technology clarifies or complicates the learning task.

The findings are relevant to vocational education because students are expected to use mathematics in conjunction with technical reasoning and digital tools. Vocational education is most effective when knowledge is connected to purposeful activity and occupational practice rather than taught as isolated theory (Billett, 2011). Mathematical logic is particularly relevant to vocational pathways involving computing, networking, electronics, data processing, and other fields in which conditional reasoning and algorithmic structure are routine. Personalized practice may therefore support not only immediate test performance but also the development of habits of checking conditions, tracing consequences, and identifying invalid reasoning. This connection should be examined directly in future studies through authentic vocational problem contexts rather than assumed from classroom test gains.

The present findings also align with large-scale evidence on adaptive mathematics technology. Cognitive Tutor Algebra research has shown that technology-supported personalized curricula can improve outcomes, although benefits may depend on implementation time and instructional fidelity (Pane et al., 2014). Personalized contexts within adaptive algebra systems have likewise been shown to improve performance for some learners by connecting tasks to students' interests (Walkington, 2013). Online mathematics environments that combine timely feedback with teacher access to learner information can also improve achievement (Roschelle et al., 2016). Taken together, these studies suggest that effective technology-enhanced mathematics learning typically involves an interaction among adaptive tools, carefully designed tasks, and teacher mediation rather than a simple substitution of software for instruction.

Recent reviews further caution against interpreting AI as a uniformly effective intervention. AIED outcomes depend on data quality, transparency, teacher competence, access to infrastructure, and the extent to which systems support rather than displace productive reasoning (Chiu et al., 2023; Crompton & Burke, 2023; Zhai et al., 2021). Mathematics-specific reviews published in 2026 similarly report that AI-driven personalization can improve problem solving and conceptual understanding but may also create over-reliance on hints, misaligned adaptivity, or reduced independent reasoning if support is excessive (Eti et al., 2026; Molemane et al., 2026). For this reason, AI-based modules should include opportunities for students to explain reasoning, attempt problems independently, and compare alternative solution strategies rather than simply follow system-generated prompts.

For teachers, the implication is that AI-supported modules should be treated as diagnostic and instructional partners. Teachers remain responsible for selecting mathematically meaningful tasks, interpreting learner data, checking whether feedback is conceptually correct, and deciding when

human explanation is needed. Recent Indonesian evidence from vocational teachers indicates generally positive perceptions of AI integration, but also highlights the need for professional development and adequate infrastructure (Yuliansah et al., 2026). Similarly, recent Indonesian and international reviews of adaptive mathematics learning emphasize teacher readiness as a condition for sustainable implementation (Maryati et al., 2026; Nurhayati et al., 2026). Professional development should therefore address not only technical operation but also AI literacy, data ethics, pedagogical decision making, and ways to prevent overdependence on automated guidance.

Successful implementation in vocational schools also requires attention to feasibility and equity. Adaptive modules depend on stable device access, connectivity, technical maintenance, and sufficient teacher time to interpret learner information. If these conditions are uneven, the same technology that supports personalization can also widen differences between students who can access the system consistently and those who cannot. Reviews of AI in education repeatedly identify infrastructure, digital competence, transparency, and data governance as conditions that shape educational value (Chiu et al., 2023; Zhai et al., 2021). In the Indonesian vocational context, teacher perspectives likewise suggest that positive attitudes toward AI need to be matched by structured professional development and institutional support (Yuliansah et al., 2026). For school-level adoption, a practical strategy is therefore to begin with clearly bounded learning objectives, verify the accuracy of automated feedback, monitor student dependence on hints, and compare digital learning evidence with teacher observation. The module should also make its adaptation logic understandable enough for teachers to identify when the system is recommending inappropriate difficulty or repeating unproductive practice. Such human oversight is particularly important in mathematics, where an automatically generated explanation may appear plausible while containing a conceptual error. Responsible integration consequently requires a balance among personalization, transparency, pedagogical control, and student independence.

Several limitations should frame interpretation of the study. First, the sample was limited to 60 students in one public vocational high school, so the findings cannot be generalized to all vocational learners. Second, the use of intact classes and purposive sampling means that unmeasured group differences may remain despite similar pretest means. Third, the intervention lasted only six meetings, which is sufficient to observe short-term improvement but not retention or transfer. Fourth, the available dataset summary does not provide exact *t* statistics, confidence intervals, or standardized effect sizes, limiting the precision of statistical interpretation. Fifth, the study evaluates the AI-supported module as a package and does not isolate which component—adaptive sequencing, automatic feedback, or error-pattern analysis—produced the largest contribution. Future research should use larger multi-school samples, longer follow-up periods, implementation-fidelity measures, and designs that compare specific AI functions.

Future studies should also broaden the outcome measures. Mathematical logic learning involves more than final scores; it includes explanation quality, transfer to unfamiliar problems, metacognitive monitoring, and the ability to detect invalid arguments. Research could combine achievement tests with process data, think-aloud protocols, log-file analysis, and delayed posttests. The use of AI-generated learner data should follow transparent privacy and governance procedures. Finally, experimental comparisons should report exact test statistics, confidence intervals, and standardized effect sizes so that results can be integrated into future meta-analyses. These steps would strengthen the evidence needed to determine when and for whom AI-supported mathematics modules are most effective.

Conclusion

This study found that Grade 10 vocational high-school students who learned mathematical logic with an AI-supported deep-learning module showed greater improvement than students who received regular instruction. The experimental class increased from a mean pretest score of 56.42 to a posttest mean of 82.15 and obtained an N-gain of 0.62, compared with an N-gain of 0.38 in the control class. Using the conventional Hake categories, both gains are classified as medium, but the experimental improvement was clearly larger and the reported between-group comparison was statistically significant ($p < .05$). These results support the short-term effectiveness of the module in the context studied.

The educational contribution of the study lies in demonstrating how AI-supported adaptivity, feedback, and error analysis can be integrated into a structured mathematics module for vocational learners. The findings are consistent with broader evidence on intelligent tutoring and adaptive mathematics learning, while extending attention to mathematical logic and the vocational secondary-school context. At the same time, the results should be interpreted cautiously because the study involved one school, intact classes, a short intervention period, and limited statistical reporting. AI should therefore be viewed as a support for sound mathematics pedagogy rather than as an autonomous substitute for teachers. Future research should test the module across multiple vocational programs, examine retention and transfer, compare individual AI functions, report standardized effect sizes, and investigate how teacher mediation and student AI literacy influence learning outcomes.

Acknowledgements

The authors gratefully acknowledge the principal, mathematics teachers, and Grade 10 students of the participating public vocational high school for their cooperation during the implementation of the study. Appreciation is also extended to the experts who contributed to the review of the learning module and research instrument, and to the Doctoral Study Program in Education, Faculty of Teacher Training and Education (FKIP), University of Bengkulu, for the academic environment and scholarly support that facilitated the completion of this research. The authors also thank colleagues who provided constructive comments on the development of the manuscript.

References

- Azwar, S. (2020). Reliabilitas dan validitas (5th ed.). Pustaka Pelajar.
- Billett, S. (2011). Vocational education: Purposes, traditions and prospects. Springer. <https://doi.org/10.1007/978-94-007-1954-5>
- British Educational Research Association. (2018). Ethical guidelines for educational research (4th ed.). BERA.
- Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 1, 100002. <https://doi.org/10.1016/j.caeai.2020.100002>
- Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. <https://doi.org/10.1016/j.caeai.2022.100118>
- Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). SAGE.
- Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field.

- International Journal of Educational Technology in Higher Education, 20, 22. <https://doi.org/10.1186/s41239-023-00392-8>
- Eti, N., Mosia, M., & Egara, F. O. (2026). The role of AI-driven personalised learning in enhancing mathematics problem-solving skills: A systematic review. *Frontiers in Computer Science*, 8, 1813431. <https://doi.org/10.3389/fcomp.2026.1813431>
- Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). SAGE.
- Guan, C., Mou, J., & Jiang, Z. (2020). Artificial intelligence innovation in education: A twenty-year data-driven historical analysis. *International Journal of Innovation Studies*, 4(4), 134–147. <https://doi.org/10.1016/j.ijis.2020.09.001>
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74. <https://doi.org/10.1119/1.18809>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Heffernan, N. T., Ostrow, K. S., Kelly, K., Selent, D., Van Inwegen, E. G., Xiong, X., & Williams, J. J. (2016). The future of adaptive learning: Does the crowd hold the key? *International Journal of Artificial Intelligence in Education*, 26(2), 615–644. <https://doi.org/10.1007/s40593-016-0094-z>
- Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S. I., & Reiss, K. M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. *Computers & Education*, 153, 103897. <https://doi.org/10.1016/j.compedu.2020.103897>
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Hwang, G.-J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 1, 100001. <https://doi.org/10.1016/j.caeai.2020.100001>
- Hwang, S. (2022). Examining the effects of artificial intelligence on elementary students' mathematics achievement: A meta-analysis. *Sustainability*, 14(20), 13185. <https://doi.org/10.3390/su142013185>
- 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. <https://doi.org/10.1016/j.caeai.2021.100017>
- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. National Academy Press.
- Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research*, 86(1), 42–78. <https://doi.org/10.3102/0034654315581420>
- Lithner, J. (2008). A research framework for creative and imitative reasoning. *Educational Studies in Mathematics*, 67(3), 255–276. <https://doi.org/10.1007/s10649-007-9104-2>
- Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, 106(4), 901–918. <https://doi.org/10.1037/a0037123>
- Maryati, I., Harun, M., Gumilar, S., & Rahayu, A. P. (2026). Adaptive learning and generative AI in mathematics teacher education: A systematic review. *International Journal of Review in Mathematics Education*, 1(1), 24–51. <https://doi.org/10.23917/ijrime.15936>

- Mayer, R. E. (2021). *Multimedia learning* (3rd ed.). Cambridge University Press.
<https://doi.org/10.1017/9781316941355>
- Molemane, N., Mosia, M., & Egara, F. O. (2026). Effectiveness of AI-based tutoring and assessment systems in mathematics education: A systematic review. *Discover Education*. Advance online publication. <https://doi.org/10.1007/s44217-026-01872-5>
- Nurhayati, S., Sudirman, S., Kartono, K., & Rodríguez-Nieto, C. A. (2026). Adaptive learning systems in mathematics education: An integrated bibliometric mapping and systematic literature review. *Kalamatika: Jurnal Pendidikan Matematika*, 11(1), 20–46.
<https://doi.org/10.22236/KALAMATIKA.vol11no1.2026pp20-46>
- Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, 2, 100020.
<https://doi.org/10.1016/j.caeai.2021.100020>
- Pane, J. F., Griffin, B. A., McCaffrey, D. F., & Karam, R. (2014). Effectiveness of Cognitive Tutor Algebra I at scale. *Educational Evaluation and Policy Analysis*, 36(2), 127–144.
<https://doi.org/10.3102/0162373713507480>
- Ran, H., Kasli, M., & Secada, W. G. (2021). A meta-analysis on computer technology intervention effects on mathematics achievement for low-performing students in K-12 classrooms. *Journal of Educational Computing Research*, 59(1), 119–153.
<https://doi.org/10.1177/0735633120952063>
- Roschelle, J., Feng, M., Murphy, R. F., & Mason, C. A. (2016). Online mathematics homework increases student achievement. *AERA Open*, 2(4), 1–12.
<https://doi.org/10.1177/2332858416673968>
- Schoenfeld, A. H. (1985). *Mathematical problem solving*. Academic Press.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189.
<https://doi.org/10.3102/0034654307313795>
- Son, T. (2024). Intelligent tutoring systems in mathematics education: A systematic literature review using the substitution, augmentation, modification, redefinition model. *Computers*, 13(10), 270. <https://doi.org/10.3390/computers13100270>
- Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on K-12 students' mathematical learning. *Journal of Educational Psychology*, 105(4), 970–987. <https://doi.org/10.1037/a0032447>
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31, 261–292.
<https://doi.org/10.1007/s10648-019-09465-5>
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273–1296.
<https://doi.org/10.1007/s11165-016-9602-2>
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221.
<https://doi.org/10.1080/00461520.2011.611369>
- Walkington, C. (2013). Using adaptive learning technologies to personalize instruction to student interests: The impact of relevant contexts on performance and learning outcomes. *Journal of Educational Psychology*, 105(4), 932–945. <https://doi.org/10.1037/a0031882>
- Wang, X., Huang, R. T., Sommer, M., Pei, B., Shidfar, P., Rehman, M. S., Ritzhaupt, A. D., & Martin, F.

- (2024). The efficacy of artificial intelligence-enabled adaptive learning systems from 2010 to 2022 on learner outcomes: A meta-analysis. *Journal of Educational Computing Research*, 62(6), 1348–1382. <https://doi.org/10.1177/07356331241240459>
- Yuliansah, Y., Sholikhah, M., & Sutirman, S. (2026). Exploring artificial intelligence in vocational learning: Teachers' perspective from Indonesia. *IAES International Journal of Artificial Intelligence*, 15(3), 2009–2023. <https://doi.org/10.11591/ijai.v15.i3.pp2009-2023>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & 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, 39. <https://doi.org/10.1186/s41239-019-0171-0>
- Zhai, X., Chu, X., Chai, C. S., Jong, M. S.-Y., Istenič, A., Spector, M., Liu, J.-B., Yuan, J., & Li, Y. (2021). A review of artificial intelligence in education from 2010 to 2020. *Complexity*, 2021, 8812542. <https://doi.org/10.1155/2021/8812542>