IJGIE Scientific Journal
International Journal of Graduate of Islamic Education

ARTIFICIAL INTELLIGENCE-BASED TEACHER TRAINING IN THE CONTEXT OF COASTAL SCHOOLS: EVALUATION USING THE KIRKPATRICK MODEL

Kamsuri[1]; Dody Irawan[2]; Encik Abdulhajar[3]; Tety Kurmalasari[4]; Zaitun[5]; Nevrita[6]
[1,2,3,4,5,6]*Program magister pedagogi, Universitas Maritim Raja Ali Haji, Tanjungpinang, Kepualauan Riau, Indonesia*
Corresponding E-mail: kamsuri59@admin.sma.belajar.id

Abstract

The development of artificial intelligence (AI) provides new opportunities for improving teachers' professional competencies, particularly in the development of digital learning tools. However, the implementation of AI-based training in coastal schools still faces obstacles, mainly related to infrastructure limitations and digital literacy. This study aims to evaluate the effectiveness of the Teacher Capacity Building Programme in the Use of AI for the development of learning tools at SMA Negeri 2 Bintan Pesisir, Riau Islands Province. The evaluation was conducted using Kirkpatrick's four-level model, which includes reaction, learning, behaviour, and results. This study used a mixed methods approach with data collection through questionnaires, pretest–posttest tests, and observation. The results showed that the training programme obtained an overall average score of 84.3, which is in the good to very good category. Specifically, the participants' reaction level scored 86.8, learning level 84.4, behavioural change 80.3, and final results 85.7. Correlation analysis between levels showed a coefficient of 0.87, indicating a strong linear relationship between evaluation components. These findings indicate that AI-based training contributes positively to improving teachers' ability to design digital learning tools, although its implementation is still influenced by contextual factors in schools. Theoretically, this study reinforces the relevance of the Kirkpatrick model in evaluating digital training. Practically, the results of this study can be used as a reference in designing technology-based teacher training programmes, especially in coastal areas and regions with digital limitations.

Keywords: AI, Teacher Professional Development; Kirkpatrick Evaluation Model; Digital Learning Materials; Coastal School Context

A. Introduction

The development of digital technology in the last two decades has brought fundamental changes to educational practices, particularly in how teachers design, implement, and evaluate learning. Digital technology not only expands access to learning resources but also requires teachers to have adequate pedagogical and technical skills to utilise the technology effectively (Napitupulu et al., 2025) . One of the most notable developments is the advent of artificial intelligence (AI), which is beginning to be utilised in various aspects of education, including learning administration management and support in the design of teaching tools, as well as strengthening teachers' pedagogical readiness to face the demands of the 21st century (Napitupulu et al., 2025) .

In the context of teacher professional development, improving digital competence through training integrated with pedagogical practice has been proven to correlate

positively with improved teaching skills and professional well-being, as effective training must simultaneously cover technology, pedagogy, and content reinforcement (Shi et al., 2025) . However, the use of AI in education cannot be separated from the social, geographical, and human resource readiness context of each educational unit. Recent empirical research highlights that infrastructure limitations, digital access gaps, and a lack of intensive training are major obstacles to the integration of technology into teaching and learning practices, especially in economically and geographically disadvantaged schools (Yulaikhah et al., 2026) . This gap indicates that the successful utilisation of educational technology is not only determined by its sophistication but also by the readiness and capacity of teachers as the primary users and adequate institutional support (OECD, 2025) .

Various national education policies have encouraged digital transformation through strengthening teacher competencies, including through technology-based training programmes. However, most training programmes still focus on technical introduction to the use of applications without comprehensive evaluation of their impact on changes in teacher behaviour and the quality of learning in the classroom (Anisah et al., 2025) . A number of studies show that although technology training can improve teachers' basic skills in digital technology, sustainable changes in learning practices are highly dependent on school environment support, responsive policies, and continuous competency development (Shi et al., 2025; Yulaikhah et al., 2026) .

In the context of coastal schools, these challenges become even more complex due to a combination of infrastructure limitations and social factors that influence the intensity of technology utilisation. Therefore, teacher capacity building programmes in AI utilisation need to be systematically evaluated to ensure that the training provided is truly relevant, applicable, and has an impact on teaching practices. The evaluation of teacher training programmes should not only assess participant satisfaction or knowledge improvement, but also behavioural changes and learning outcomes . Kirkpatrick's Four-Level Evaluation Model offers a comprehensive evaluation framework that assesses participant reactions, learning that has taken place, behavioural changes in the workplace, and the impact of training outcomes on the organisation.

Based on these conditions, there is a research gap related to the lack of empirical studies that comprehensively evaluate the effectiveness of AI utilisation training programmes for teachers, taking into account the context of regional limitations. This study offers novelty by comprehensively applying the Kirkpatrick Model to evaluate teacher capacity building programmes in AI utilisation in coastal schools. The purpose of this study is to analyse the effectiveness of training programmes in terms of reaction, learning, behavioural change, and results, as well as to identify contextual factors that influence their implementation.

B. Method

This study uses a mixed methods approach with an explanatory sequential design, in which quantitative data is collected and analysed first, then explored further through qualitative data. This approach was chosen because it allows researchers to gain a more comprehensive understanding of a programme's effectiveness, not only from numerical achievements but also from the context, experiences, and dynamics of implementation in the field. The explanatory sequential design is considered effective for evaluative research

because qualitative data can be used to explain and reinforce quantitative findings obtained in the initial stage ((Creswell & Plano Clark, 2023) .

The research was conducted at a state senior secondary school located in a coastal area, which contextually has limited infrastructure and access to technology. The research subjects consisted of all teachers who participated in a series of training programmes on the use of Artificial Intelligence (AI). All participants were included as research subjects because their number was limited and they represented the target population of the programme. The use of total sampling in the context of small-scale training was considered appropriate to obtain a complete evaluative picture without losing data variation between participants (Etikan & Bala, 2021).

The research instrument was developed based on the four levels of evaluation in the Kirkpatrick Model, which include reaction, learning, behaviour, and results. This model was chosen because it is capable of evaluating the effectiveness of training in stages, starting from the initial response of participants to the impact of training outcomes on work practices and performance quality. In the context of technology-based teacher training, the Kirkpatrick model is considered relevant because it allows for assessment not only of cognitive aspects, but also of behavioural changes and implementation outcomes in the school environment (Kirkpatrick & Kirkpatrick, 2021) .

The research instruments were developed based on indicators from each level of evaluation in the Kirkpatrick Model, which includes reaction, learning, behaviour, and results. The use of the Kirkpatrick Model in the development of training evaluation instruments is considered relevant because this model is able to provide a comprehensive picture of the effectiveness of the programme, starting from the initial response of participants to the real impact in the work context (Kirkpatrick & Kirkpatrick, 2021). The content validity of the instrument was consulted with three experts in education and learning technology to ensure the suitability of the indicators with the training objectives and the implementation context in coastal schools. This expert validation process is important because content validity is a key prerequisite for the evaluation instrument to truly represent the construct being measured and have practical relevance to the research context (Taherdoost, 2021) . In addition, the involvement of cross-disciplinary experts strengthens the conceptual accuracy of the instrument, particularly in integrating pedagogical and technological aspects in AI-based training evaluation. This approach is in line with recent research recommendations emphasising that the evaluation of technology-based teacher professional development programmes needs to be supported by systematically validated instruments so that the evaluation results can be used as a basis for decision-making and the development of sustainable education policies (Darling-Hammond et al., 2020).

Table 1. Kirkpatrick Evaluation Focus

Evaluation Level	Evaluation Focus
Level 1: Reaction	Teachers' perceptions and satisfaction with the training programme (content, methods, facilitators, and relevance)
Level 2: Learning	Improvement in knowledge and skills in using AI to create learning tools
Level 3 Behaviour	Changes in teachers' behaviour in implementing AI in the classroom
Level 4 outcomes	Impact of the programme on the quality of learning tools and the effectiveness of learning in the classroom

Data at the reaction level was collected using a Likert scale questionnaire to measure participants' perceptions and satisfaction with the material, methods, facilitators, and relevance of the training. The learning level was measured through pre- and post-tests and assessments of practical tasks in developing AI-based learning tools, which aimed to objectively capture participants' increased knowledge and skills. The behavioural level was examined through classroom observations and analysis of the learning tools used by teachers after the training, as behavioural change is considered an important indicator of the transfer of training outcomes into real-world practice (Blume et al., 2022) . Meanwhile, the outcome level was analysed through an assessment of the quality of learning tools and interviews with school management to see the impact of the training at the organisational level.

Quantitative data were analysed using descriptive statistics and N-Gain calculations to identify improvements in teacher learning outcomes before and after training. The relationship between evaluation levels was analysed using Pearson's correlation to examine the consistency of training outcomes from the reaction level to the results level. Qualitative data were analysed thematically to identify patterns of implementation, constraints, and teachers' perceptions of the use of AI in learning. The integration of quantitative and qualitative data was carried out at the interpretation stage to produce stronger and more contextual conclusions, as recommended in mixed methods-based evaluation research (Guetterman & Fetters, 2021) .

The research instruments were developed based on indicators at each level of Kirkpatrick's evaluation and were consulted with education and learning technology experts to ensure content validity. Instrument validity through expert judgement is considered important in evaluative research so that the indicators used truly represent the constructs being measured, especially in the context of multidimensional technology-based training (Boateng & Neilands, 2024).

C. Findings and Discussion

The evaluation design in this study was designed to answer the hypothesis that teacher capacity building training programmes can significantly improve teachers' ability to utilise AI for developing learning tools. Systematic evaluation results across the four levels of Kirkpatrick's model show that this hypothesis can be accepted, as the programme has a measurable positive impact on each dimension. The main quantitative data generated from this study is summarised in the following table:

Table 2. Summary of Programme Evaluation Results

Evaluation Level	Key Indicators	Average Score	Category
Level 1 - Reaction	Satisfaction with the material and facilitator	86	Very good
Level 2 - Learning	Knowledge gain (N-Gain 0.59)	84.4	High
Level 3 – Behaviour	Implementation of AI in the classroom	80.3	Good
Level 4 – Outcomes	Impact on the quality of teaching materials	85.7	Very good

The results of the evaluation of the Artificial Intelligence (AI)-based teacher training programme were analysed using Kirkpatrick's four-level model to obtain a comprehensive picture of the programme's effectiveness, from the initial responses of participants to the resulting impact. This evaluation design allowed researchers to not only assess quantitative achievements but also understand the dynamics of training implementation in the context of coastal schools, which have their own characteristics and limitations.

Based on Table 2, at the reaction level, an average score of 86.8 was obtained, which is in the excellent category. This achievement shows that participants responded positively to the training, especially regarding the relevance of the material, the training approach, and the facilitators' competence. This positive reaction plays an important role because, in the training evaluation framework, participant satisfaction and acceptance are prerequisites for effective and meaningful learning processes (Kirkpatrick & Kirkpatrick, 2021) . These findings indicate that the training material was in line with the actual needs of teachers, particularly in facing the demands of developing digital learning tools that are adaptive to technological developments. Conceptually, these results are in line with the view that the initial success of a training programme is greatly influenced by the suitability of the material to the work context and professional needs of the participants (Darling-Hammond et al., 2020) .

At the learning level, the increase in teachers' understanding of the basic concepts of AI and its application is reflected in an N-Gain value of 0.59, which is in the moderate category, with an average score of 84.4. These results indicate that the training was able to significantly improve teachers' cognitive competencies compared to the initial conditions, although it has not yet reached the optimal level. The moderate learning achievement can be understood given the complexity of AI material and the limitations of teachers' initial experience in utilising this technology. Recent research shows that mastery of AI-based technology requires a gradual and continuous process so that conceptual understanding can develop into applicable skills in learning practice (Holmes et al., 2022) . Thus, the findings of this study are consistent with the literature which emphasises that technology training is generally effective as an introductory stage, but requires follow-up for a more profound impact.

Evaluation at the behavioural level showed an average score of 80.3, which is in the good category. Observation and documentation data indicate that some teachers have begun to integrate AI into lesson planning and the development of digital teaching materials. However, this behavioural change has not occurred uniformly among all participants. This variation in implementation is influenced by a number of contextual factors, such as limited internet access, availability of supporting devices, and teaching habits that are still oriented towards conventional methods. This condition reinforces the view that increased knowledge does not automatically correlate with changes in work practices, because training transfer is greatly influenced by environmental support, organisational culture, and the readiness of educational institutions (Blume et al., 2022) .

At the outcome level, an average score of 85.7 indicates that the training programme has a positive impact on the quality of learning materials produced by teachers. Document analysis shows improvements in the structure of teaching tools, clarity in the formulation of learning objectives, and a greater variety of learner-centred learning activities. However, the direct impact on the learning process in the classroom is not yet fully optimal. This confirms that in the context of this study, AI functions as a pedagogical tool that supports teachers' creativity and efficiency, rather than replacing the professional role of teachers in managing learning. These findings are relevant to the digital education literature, which emphasises that technology only has a significant impact when it is integrated reflectively and pedagogically, rather than simply being used technically (Holmes et al., 2022; OECD, 2025).

Overall, the strong linear relationship between evaluation levels with a correlation coefficient of 0.87 shows the consistency of programme achievements from reactions to results. However, the effectiveness of the programme cannot be understood solely from quantitative scores. The dynamics of implementation in the field show that the coastal school context plays an important role in determining the extent to which training results can be internalised and applied. This is in line with recent studies that emphasise that geographical context and the availability of digital infrastructure are determining factors for the successful integration of technology in education (UNESCO, 2023) . Therefore, the results of this study are scientifically acceptable because they are in line with training evaluation theory and supported by empirical data, while also providing a contextual contribution to the literature on technology-based teacher training in areas with limited infrastructure.

The implications of this research indicate that AI training programmes for teachers need to be designed as a continuous process accompanied by institutional guidance and support. The limitations of this research lie in the scope of the location and the relatively short duration of the behavioural evaluation. Further research is recommended to explore the long-term impact of AI training on classroom learning practices and to involve more educational units with diverse regional characteristics, as recommended in the latest technology-based training evaluation research (Darling-Hammond et al., 2020) .

E. Conclusion

This study shows that the Artificial Intelligence (AI) utilisation training programme for teachers has a positive impact on improving knowledge and the quality of learning tools, but behavioural changes and learning outcomes in the classroom are still influenced by the limitations of the coastal school context. These findings are in line with recent studies which confirm that the success of educational technology training is not only determined by the improvement of participants' cognitive competencies, but also by the readiness of the work environment and institutional support that enables the transfer of training outcomes into learning practices (Darling-Hammond et al., 2020) . In this context, AI has been proven to assist teachers, particularly in improving work efficiency, such as in the preparation of learning tools and the management of digital learning resources. However, it cannot replace the pedagogical role of teachers as mentors, decision-makers, and facilitators of learning oriented towards the needs of students. This view is in line with the latest educational literature, which positions AI as *a pedagogical support tool* that serves to strengthen, not replace, teachers' professionalism (Holmes et al., 2022) . Therefore, AI training for teachers needs to be designed in a sustainable and contextual manner and accompanied by mentoring so that its application is not temporary or merely a response to technological trends. These findings confirm that the successful integration of AI in education is highly dependent on the balance between technological sophistication, teachers' pedagogical competence, and the actual conditions of educational units, especially in areas with limited infrastructure such as coastal schools (UNESCO, 2023) .

G. Bibliography

Anisah, A. S., Maspirah, & Putri, M. J. (2025). Rethinking Teacher Professional Development for the Digital Era: A Narrative Review. *Eduscape: Journal of Education Insight*, *3*(4), 204–217.

Blume, B. D., Ford, J. K., & Baldwin, T. T. (2022). Transfer of Training: A Meta-Analytic Review. *Journal of Management*, *48* (4), 987–1022. https://doi.org/10.1177/01492063211047557

Boateng, G. O., & Neilands, T. B. (2024). Best Practices for Developing and Validating Scales for Educational Research. *Educational Measurement: Issues and Practice*, *43* (1), 5–18. https://doi.org/10.1111/emip.12509

Creswell, J. W., & Plano Clark, V. L. (2023). *Designing and Conducting Mixed Methods Research* (3rd ed.). Sage Publications.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2020). *Effective Teacher Professional Development*.

Etikan, I., & Bala, K. (2021). Sampling and Sampling Methods. *Biometrics & Biostatistics International Journal*, *10*(1), 7–11.

Guetterman, T. C., & Fetters, M. D. (2021). Integration in Mixed Methods Research. *Journal of Mixed Methods Research*, *15* (3), 293–306. https://doi.org/10.1177/1558689820988594

Holmes, W., Bialik, M., & Fadel, C. (2022). Artificial Intelligence in Education: Promise and Implications for Teaching and Learning. *Educational Technology Research and*

Development, *70* (5), 2401–2420. https://doi.org/10.1007/s11423-022-10124-9

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2021). *Four Levels of Training Evaluation*. ATD Press.

Napitupulu, M. H., Muddin, A., Bagiya, B., Diana, S., & Rosyidah, N. S. (2025). Teacher Professional Development in the Digital Age: Strategies for Integrating Technology and Pedagogy. *International Journal for Science Review*, *2*(10).

OECD. (2025). *Preparing Teachers for Digital Education*.

Shi, Y. R., Sin, K. F. K., & Wang, Y. Q. (2025). Teacher professional development of digital pedagogy for inclusive education in the post-pandemic era. *Teaching and Teacher Education*, *168* , 105230. https://doi.org/10.1016/j.tate.2025.105230

Taherdoost, H. (2021). Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research. *International Journal of Academic Research in Management*, *10* (1), 1–6. https://doi.org/10.2139/ssrn.3205040

UNESCO. (2023). *Technology in Education: A Tool on Whose Terms?*

Yulaikhah, S., Suhendra, R. R., & Hidayat, N. (2026). Systematic Literature Review of Barriers to Teachers' Professional and Digital Competence in Developing Contexts (2020–2025). *Journal of Informatics, Education and Management*, *8*(1), 540–552.