PROFESI PENDIDIKAN DASAR
Volume 13, Issue 1, April 2026, pp. 90-104
p-ISSN: 2406-8012, e-ISSN: 2503-3530
https://journals2.ums.ac.id/index.php/ppd

AI-Supported Deep Learning Model for Meaningful Character Education in Indonesian Elementary School Learning

Lilik Binti Mirnawati[1*], Aswin Rosadi[1], Apostolis Xenakis[2], Ummi Hani Abu Hassan[3]
[1]Universitas Muhammadiyah Surabaya, Indonesia
[2]University of Thessaly, Greece
[3]Hankuk University of Foreign Studies, Republic of Korea
*Corresponding Author's email: liIikbintimirnawati@um-surabaya.ac.id

How to cite: Mirnawati, L. B., Rosadi, A., Xenakis, A., & Hassan, U. H. A. (2026). AI-Supported Deep Learning Model for Meaningful Character Education in Indonesian Elementary School Learning. *Profesi Pendidikan Dasar*, *13*(1), 90–104. https://doi.org/10.23917/ppd.v13i1.14516

Keywords:
artificial intelligence; meaningful learning; character education; Indonesian language learning

Article History:
Submitted: 2025-12-10
Revised: 2026-02-16
Accepted: 2026-03-15

Abstract

Elementary literacy performance in Indonesia remains below international standards, indicating the need for instructional innovation within core subjects. This study developed and evaluated an AI-supported deep learning model implemented in Bahasa Indonesia learning to enhance meaningful learning and character education in elementary schools. The research employed a Research and Development design using a sequential explanatory mixed-methods approach. The development process included needs analysis, expert validation, revision, pilot testing, and field trials involving 60 students and two teachers. Quantitative data were collected through meaningful literacy tests, character observation rubrics, learning analytics logs, and perception questionnaires, while qualitative data were obtained from interviews and classroom observations. Paired-samples t-tests showed significant improvements in conceptual understanding, critical thinking, and knowledge application. There was a notable enhancement in character indicators, particularly in the domains of empathy and responsibility. An examination of the analytics revealed a marked increase in student engagement over a period of four weeks. The integration of personalisation and analytics within the teaching of Bahasa Indonesia has been demonstrated to enhance cognitive and personal development in elementary education.

INTRODUCTION

Background of the Study

In the 21st century, education ought to embrace practices that not only enhance factual knowledge but also nurture critical thinking, creativity, collaboration, digital literacy and character (Holmes & Tuomi, 2022). Artificial Intelligence (AI) has increasingly been recognized as a promising innovation capable of supporting these multidimensional competencies through adaptive learning pathways, personalized feedback, and real-time learning analytics (Rojas & Chiappe, 2024; Sajja et al., 2024). In elementary education, the integration of AI is particularly strategic because this level represents a foundational stage for cognitive growth and value internalization. However, effective AI implementation requires pedagogical alignment, teacher readiness, and contextual adaptation within specific subject areas (Kim & Kwon, 2023).

Within the Indonesian context, the urgency of instructional innovation is strongly reflected in students' literacy performance. The PISA 2022 results indicate that Indonesia's reading literacy score (359) remains significantly below the OECD average (476), suggesting persistent challenges in comprehension, critical interpretation, and contextual application of texts. These issues are directly connected to the teaching of Bahasa Indonesia as a core subject responsible for developing reading, writing, and interpretative competencies in elementary schools. Despite its central role, Bahasa Indonesia instruction often remains teacher-centered and text-reproduction oriented, limiting opportunities for deep processing, reflection, and character development. Therefore, integrating AI-supported deep learning strategies into Bahasa Indonesia learning becomes a relevant and urgent approach to strengthening meaningful literacy and character education simultaneously.

Problem of the Study

Despite the rapid advancement of Artificial Intelligence (AI) in global education, its application in Indonesian elementary schools remains limited, particularly within Bahasa Indonesia learning as a core subject responsible for developing literacy and interpretative competence. In practice, teachers commonly use AI merely to generate instructional materials rather than as a pedagogical tool that promotes exploration, reflection, and meaningful engagement (Efriyanti et al., 2024; Zhou & Peng, 2025). As a result, the transformative potential of AI to support deep literacy processing and character formation in Bahasa Indonesia classrooms has not been optimally realized.

In many elementary Bahasa Indonesia classes, learning remains teacher-centered, emphasizing text reproduction and procedural completion rather than critical interpretation and contextual application. Such instructional patterns reduce student agency and limit opportunities for active meaning-making (Kain et al., 2024; Santos-Meneses & Drugova, 2023). Consequently, the cognitive, affective, and social dimensions of learning character values such as empathy, responsibility, collaboration, and integrity are insufficiently integrated into classroom practice (Manganello & Baldacci, 2024; Zainuddin et al., 2025). This condition highlights a significant gap: the absence of a structured and holistic AI-supported deep learning model specifically designed to strengthen meaningful literacy and character education within Bahasa Indonesia instruction at the elementary level.

Research's State of the Art

Recent studies demonstrate that Artificial Intelligence (AI) has been widely implemented to support personalized learning, adaptive assessment, and literacy development across educational levels (Arriola-Mendoza & Valerio-Ureña, 2024; Kosmas et al., 2025). In elementary education, AI tools have been utilized to enhance reading comprehension, feedback mechanisms, and student engagement. However, most studies tend to examine isolated outcomes such as motivation, comprehension scores, or digital interaction patterns, without integrating cognitive, affective, and character dimensions within a unified pedagogical framework (Boulhrir & Hamash, 2025). Within Bahasa Indonesia learning, AI implementation remains largely instrumental, focusing on content generation or automated feedback rather than fostering deep literacy practices that promote reflection, interpretation, and value internalization.

Conceptually, AI-Driven Deep Learning refers to the integration of artificial intelligence technologies with deep learning pedagogy to facilitate adaptive, inquiry-oriented, and student-centered instructional processes. Deep learning in education emphasizes higher-order thinking, metacognitive engagement, and the ability to apply knowledge in authentic contexts (Moundridou et al., 2024; Santos-Meneses & Drugova, 2023). When supported by AI, this approach allows real-time analytics, personalized scaffolding, and dynamic adjustment of learning tasks based on student performance patterns (Rojas & Chiappe, 2024). In the context of Bahasa Indonesia learning, AI-driven deep learning can potentially strengthen meaningful literacy practices by guiding students beyond surface-level text comprehension toward critical analysis and contextual interpretation.

Meaningful learning, rooted in cognitive learning theory, occurs when new information is actively connected to prior knowledge, enabling deeper conceptual understanding rather than rote memorization (Mayer, 2024). In Bahasa Indonesia instruction, meaningful learning involves interpreting texts, constructing arguments, and applying linguistic knowledge in real-life communication contexts. Simultaneously, character development in elementary school students refers to the cultivation of values such as empathy, responsibility, collaboration, and integrity through structured learning experiences (Zainuddin et al., 2025). Despite increasing attention to AI integration and character education separately, empirical studies that combine AI-driven deep learning, meaningful literacy, and structured character development within Bahasa Indonesia classrooms remain limited, particularly in the Indonesian elementary school context.

Gap Study and Objective

In many Indonesian elementary classrooms, Bahasa Indonesia learning is still dominated by teacher-centered instructional practices. Based on preliminary observations conducted during the needs analysis stage, classroom activities frequently emphasize completing textbook exercises and reproducing textual information rather than encouraging students to interpret texts critically or connect them with real-life contexts. Consequently, students often demonstrate surface-level comprehension and limited engagement in reflective literacy practices. Previous studies have demonstrated that Artificial Intelligence has been widely integrated into education to support personalization, adaptive assessment, and student engagement (Holmes & Tuomi, 2022; Rojas & Chiappe, 2024; Sajja et al., 2024). In elementary contexts, AI has been utilized to enhance literacy instruction and classroom interaction (Kosmas et al., 2025; Zhou & Peng, 2025). However, most existing studies tend to focus on isolated outcomes such as motivation, performance prediction, or technological acceptance rather than integrating cognitive depth and character formation within a unified pedagogical framework (Boulhrir & Hamash, 2025; Fan et al., 2025). Moreover, while deep learning pedagogy emphasizes metacognitive reflection, conceptual understanding, and real-life application (Gao, 2025; Mayer, 2024), empirical research connecting AI-supported instruction with these deep learning principles in elementary Bahasa Indonesia classrooms remains limited.

In the Indonesian context, research on AI integration in primary education is still largely descriptive and focuses on teacher training or media development rather than validated instructional models (Efriyanti et al., 2024). At the same time, national curriculum reforms through the Kurikulum Merdeka and the Profil Pelajar Pancasila emphasize meaningful literacy and character education (Kemendikdasmen, 2025; Zainuddin et al., 2025). However, there is a lack of empirically tested AI-supported models that systematically align deep learning principles with the development of literacy competence and character values in Bahasa Indonesia learning. This indicates a theoretical and practical gap between AI innovation, deep learning pedagogy, and character-oriented instruction at the elementary level. Based on these gaps, this study aimed to develop and empirically evaluate an AI-supported deep learning model implemented in Bahasa Indonesia learning at the elementary level. Specifically, this research sought to: (1) design an instructional model integrating AI personalization, learning analytics, and deep learning stages; (2) validate the model through expert judgment and iterative revision; (3) examine its effectiveness in improving meaningful literacy outcomes and character development; and (4) analyze student engagement patterns through learning analytics data.

METHOD

Type and Design

This study employed a Research and Development (R&D) design integrated with a sequential explanatory mixed-methods approach to develop, validate, and evaluate an AI-supported deep learning model implemented in Bahasa Indonesia learning. The product developed was a an AI-assisted instructional model integrating Google Forms-based student responses with ChatGPT-supported analytical feedback to facilitate teacher reflection and character reinforcement to support meaningful literacy and character education. The platform was structured according to deep learning stages (engage, explore, explain, elaborate, evaluate) and embedded AI-driven response analysis to provide real-time formative feedback. The R&D process followed an iterative eight-stage cycle consisting of: (1) needs analysis, (2) initial model design, (3) expert validation, (4) revision, (5) small-group trial, (6) field testing, (7) evaluation, and (8) final refinement. This structure was adapted from contemporary design-based and developmental research frameworks that emphasize iterative validation, contextual testing, and empirical refinement in authentic classroom settings (Hoadley & Campos, 2022; Simard et al., 2025). The model was selected because it allows systematic product improvement through formative evaluation and real-world implementation, making it more suitable than linear development models for educational innovation research.

The small-group trial was conducted to examine initial usability, technical functionality, and student interaction patterns in a controlled classroom setting. This stage focused on identifying operational weaknesses and interface issues before broader implementation. In contrast, the field test aimed to evaluate the effectiveness of the revised model in authentic learning environments, measuring its impact on students' meaningful learning outcomes, character development, and engagement patterns. This staged evaluation ensured gradual refinement and methodological rigor. A sequential explanatory mixed-methods strategy was applied, beginning with quantitative data collection through learning outcome tests, AI activity logs, and perception questionnaires, followed by qualitative procedures including interviews and classroom observations to interpret and contextualize statistical findings (Draucker et al., 2020; Soicher et al., 2024).

Data and Data Sources

Data for this study were obtained from students, teachers, and experts involved in each R&D stage. The research was conducted in two elementary schools in Surabaya, selected based on technological readiness (availability of laptops/tablets and reliable internet access) and teacher willingness to implement the AI-based model. The participant groups included 1) Needs analysis, consisting of two principals, six teachers, and eight students. 2) Expert validation, consisting of three experts (educational technology, curriculum, and child development psychology). 3) Small group pilot, consisting of one fourth-grade teacher and twelve students. 3) Field trial, consisting of 60 fourth-grade students and two teachers. These participants provided cognitive, affective, technological, and perceptual data to evaluate the feasibility, effectiveness, and acceptability of the AI-based instructional model.

Data Collection Technique

Data were collected using multiple techniques aligned with the embedded mixed-methods design, integrating quantitative and qualitative evidence to ensure methodological rigor. Table 1 presents the research instruments used to collect cognitive, affective, technological, validation, and perceptual data across the R&D stages.

Table 1. Research Instruments and Achievement Indicators

Instrument Type	Description	Measurement Objective	Source
Cognitive	20-item meaningful learning test (10 MCQ, 5 essays, 5 contextual case items)	Conceptual understanding, application, and critical thinking	(Dubinsky & Hamid, 2024; González-Erena et al., 2025)
Affective	Observation rubric, 3 dimensions (cooperation, responsibility, empathy)	Student character development	(Borna et al., 2024)
Technology	AI activity logs: frequency, duration, progress	Student engagement and performance patterns	(Fan et al., 2025; Wang, 2024)
Expert Validation	Expert evaluation sheet	Content accuracy, instructional design, and technical feasibility	(Kim & Kwon, 2023)
Response Questionnaire	Teacher–student perception survey	Acceptance, usability, and future improvement	(Kosmas et al., 2025)

All cognitive instruments achieved high internal consistency reliability (Cronbach's α = 0.85). The affective observation rubric demonstrated strong inter-rater reliability (Cohen's κ = 0.82). Expert validation items met acceptable content validity thresholds based on Aiken's V coefficient. The expert validation instrument employed a five-point Likert scale ranging from 1 (very unfeasible) to 5 (highly feasible) to evaluate content accuracy, instructional design, language clarity, and technical suitability. Figure 1 shows the R&D stages adapted from the Dick and Carey model, covering systematic steps from identifying objectives, conducting instructional analysis, developing strategies and materials, to formative and summative evaluation.

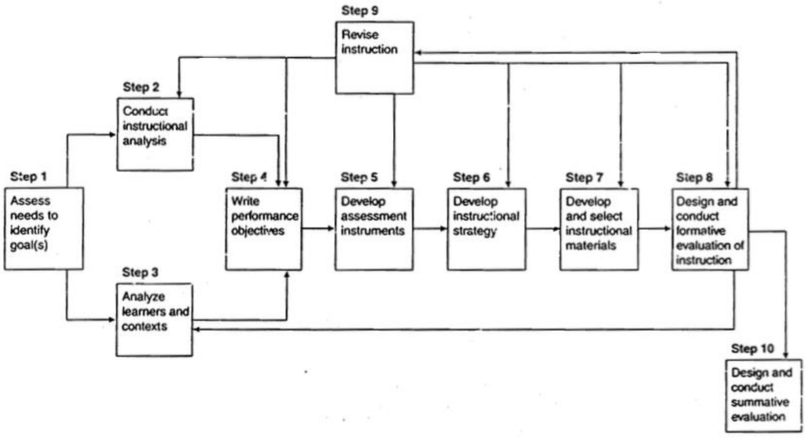

Figure 1. R&D Stages of the AI-based Instructional Model

Quantitative data were collected through several complementary instruments during the implementation stages. These included pre-test and post-test cognitive assessments administered before and after the implementation of the AI-based instructional model in the field trial phase. In addition, AI system-generated activity logs were utilized to automatically record student interaction data, including frequency of access, duration of use, task completion rates, and learning progression patterns. Teacher and student response questionnaires were also administered at the conclusion of both the small-group pilot and field trial stages. The response questionnaires employed a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The student questionnaire consisted of 15 items measuring perceived ease of use, engagement and motivation, clarity of AI-generated feedback, and overall learning satisfaction. Meanwhile, the teacher questionnaire comprised 18 items assessing instructional usability, alignment with curriculum objectives, classroom management feasibility, perceived impact on student learning outcomes, and sustainability for future implementation. Prior to full-scale administration, all questionnaire items were subjected to expert review and pilot testing to ensure clarity, relevance, and content validity.

Qualitative data were obtained through multiple methods to capture in-depth contextual insights. Semi-structured interviews were conducted with teachers and selected students after the completion of the field trial to explore instructional experiences, perceived strengths and limitations of the model, and recommendations for refinement. Classroom observations were systematically carried out using a structured rubric to document behavioural engagement and affective indicators during instructional activities. In addition, Focus Group Discussions (FGDs) with participating teachers were organized to examine pedagogical integration processes and contextual implementation constraints.

To enhance credibility and interpretative validity, triangulation was implemented at three levels. First, data source triangulation was conducted by comparing learning outcomes (pre-test and post-test scores) with AI activity logs and classroom observation data to examine consistency between measured academic achievement and observed behavioural engagement. Second, methodological triangulation involved cross-validating quantitative findings—such as test score improvements, log analytics, and questionnaire responses—with qualitative interview and FGD data to interpret underlying instructional mechanisms and contextual influences. Third, investigator triangulation was applied in the observation process, where two independent observers rated student behaviours and resolved discrepancies through consensus discussions to ensure reliability. Data integration occurred during the interpretation phase by comparing convergent and divergent patterns across datasets. For instance, improvements in post-test scores were analysed alongside AI engagement metrics and student interview responses to determine whether performance gains corresponded with meaningful interaction patterns and perceived instructional value.

Data Analysis

Data analysis was conducted separately for quantitative and qualitative datasets before being integrated at the interpretation stage. Quantitative data were analysed using both descriptive and inferential statistics with a significance level of $\alpha = 0.05$. All statistical analyses were performed using SPSS version 28. Pre-test and post-test scores from the field trial were first analysed descriptively using mean and standard deviation. Before hypothesis testing, the normality assumption was examined using the Shapiro–Wilk test due to the sample size being below 100 participants. To determine whether there was a statistically significant difference between pre-test and post-test scores, descriptive comparisons and normalized gain analysis were conducted to examine the magnitude of improvement. This test was selected because the same participants were measured before and after the intervention (one-group pre-test–post-test design). In addition to statistical significance testing, normalized gain (N-Gain) scores were calculated to determine the magnitude of learning improvement (Moore et al., 2021) using the following formula:

$$N - Gain = \frac{(Post\ test - Pretest)}{(Maximum\ score - Pre\ test)}$$

The N-Gain categories were interpreted as presented in Table 2 below:

Table 2. N-Gain Value Category

N-Gain Value	Category
$g \geq 0.70$	High
$0.30 \leq g < 0.70$	Medium
$g < 0.30$	Low

The N-Gain analysis provided a practical interpretation of improvement levels beyond statistical significance. Student character development—covering cooperation, responsibility, and empathy—was assessed using structured observation rubrics scored on a four-point scale (1 = very low; 4 = very high). Descriptive statistics (mean and standard deviation) were calculated for each character indicator before and after the intervention. To examine whether the changes were statistically significant, paired-samples t-tests were conducted for each character dimension. Additionally, a change index (Δ) was calculated to describe the magnitude of development:

$$\Delta = \bar{X}_{post} - \bar{X}_{pre}$$

Mean scores were interpreted using the following criteria (Table 3):

Table 3. Criteria for Interpreting Mean Scores

Mean Score Range	Category
3.26 – 4.00	Very Good
2.51 – 3.25	Good
1.76 – 2.50	Fair
1.00 – 1.75	Low

This approach enabled both statistical verification and descriptive interpretation of character growth. Expert validation and response questionnaire data were analyzed using mean score calculations based on a five-point Likert scale (1 = very poor; 5 = excellent). The feasibility and practicality levels were interpreted using the following criteria (Table 4):

Table 4. Criteria Feasibility and Practicality instructional model

Mean Interval	Category	Interpretation
4.21 – 5.00	Very Valid / Very Practical	Highly feasible, no revision required
3.41 – 4.20	Valid / Practical	Minor revision required
2.61 – 3.40	Moderately Valid / Moderately Practical	Revision required
1.81 – 2.60	Less Valid / Less Practical	Major revision required
1.00 – 1.80	Invalid / Impractical	Not feasible

These criteria were applied to determine the feasibility of the developed instructional model and the level of user acceptance during both the small-group trial and field testing stages. These criteria were used to interpret expert validation results in determining the validity and practical feasibility of the developed instructional model. The study employed a descriptive analysis of AI-generated learning analytics data, encompassing interaction frequency, learning duration, and idea contribution, to identify engagement growth patterns across the four-week implementation period through trend comparisons. Qualitative data from interviews, observations, and FGDs were analysed using thematic analysis procedures (Braun & Clarke, 2024), including transcription, familiarization, coding, theme generation, peer verification, and narrative synthesis. Mixed-methods integration employed triangulation by comparing quantitative findings (test score improvements, character indices, engagement metrics) with qualitative themes to identify convergent and divergent patterns. This integration informed the final refinement of the AI-supported deep learning model.

RESULTS

This section presents the findings based on the four research objectives, namely: (1) the development of the AI-supported deep learning model, (2) the validation of the developed model, (3) the practicality of the model in classroom implementation, and (4) the effectiveness of the model in improving students' literacy competencies, character development, and learning engagement.

Development of the AI-supported deep learning model

The product developed in this study is an Artificial Intelligence (AI)-based learning application designed to support meaningful learning and character education for elementary school students. This application provides personalized feedback by delivering tailored responses to students' answers, helping them identify mistakes and guiding them toward a deeper understanding of Bahasa Indonesia texts. It also incorporates learning analytics that continuously monitor and record student engagement in real time, including interaction frequency, learning duration, and contribution of ideas during learning activities. Additionally, gamification elements such as points, badges, and story-based challenges are embedded to motivate students, enhance participation, and sustain interest in learning.

The instructional design of the application follows the five stages of deep learning: engage, explore, explain, elaborate, and evaluate. In its initial version, the application is delivered through a web-based interface using Google Forms, offering a simple and intuitive design that allows students to respond to questions while teachers can access AI-generated feedback and analytics to monitor student progress. At this stage, the interface does not yet feature advanced interactive elements such as voice animation, adaptive difficulty, or multimedia scaffolds, which indicates potential areas for further development to create a more engaging and responsive learning experience.

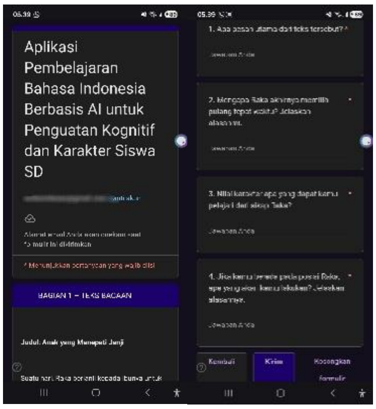

Figure 2. Display of the Web-Based AI-Based Learning Application Interface Using Google Forms

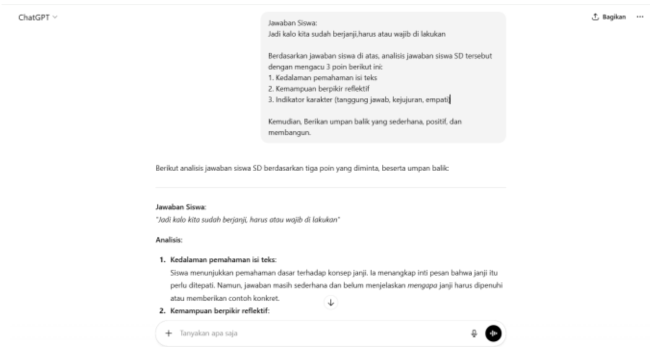

Figure 3. Analysis of Student Response Processes Using an AI System to Provide Reflective Feedback.

Figure 2 shows the initial interface of the AI-based learning application, displaying how students interact with the Google Forms interface, while Figure 3 presents an overview of the AI-driven learning analytics and feedback dashboard, illustrating how teachers can track individual student responses, engagement patterns, and conceptual understanding in real time. This description highlights both the functional structure of the application and its visual components, enabling readers to better understand how the system supports learning and instructional monitoring. Based on expert feedback, several revisions were implemented to improve usability, clarity, and instructional support of the developed product. The modifications focused on interface structure, navigation design, and multimedia enhancement to better accommodate elementary students' learning characteristics. A summary of the revisions is presented in Table 5.

Table 5. Product Design Revisions Based on Expert Validation

Component	Before Revision	After Revision
Menu Display	Simple display with dominant text	Display with clearer visual icons
Interactive Features	No voice recognition	Voice-supported interaction added
Navigation	Less clear navigation	Icons and structured navigation added
Visual & Audio	Standard visuals	Enhanced visuals and improved audio

The revisions were carried out based on expert validation to improve usability, instructional clarity, and student interaction. These refinements represent the iterative nature of the Research and Development process, ensuring that the product aligns with pedagogical needs and elementary learners' characteristics. Complete visual documentation of the revised product is provided as supplementary material and can be accessed at: https://drive.google.com/drive/folders/1gPuK0ycQ9rxKua5NI6LSNNjZysWTa--p?usp=sharing

Validity of the Developed Model

Product validation was conducted by three experts representing the aspects of content, instructional design, and technology. The validation results are presented in Table 6.

Table 6. Expert Validation Results of the Product

Feasibility Aspect	Validator 1	Validator 2	Validator 3	Average
Content Feasibility	4.2	4.5	4.3	4.33
Instructional Design	4.0	4.2	4.4	4.20
Language	4.3	4.4	4.5	4.40
Technical/Media Aspect	4.1	4.3	4.2	4.20
Total Average	**4.15**	**4.35**	**4.35**	**4.28**

The overall mean score of 4.28 indicates that the product is categorized as highly feasible, demonstrating strong validity in terms of content relevance, instructional structure, language clarity, and technical usability. Qualitative feedback from validators provided directions for refinement. Suggestions included enhancing interactivity through optional voice-response features, simplifying menu navigation for younger learners, and improving visual and audio quality to increase engagement. These inputs were used to guide subsequent product revisions.

Practicality of the Model in Classroom Implementation

User responses were analyzed across two stages of implementation in accordance with the Research and Development procedure, namely the small-group trial (limited trial) and the field trial (wider implementation). These stages were conducted to evaluate the practicality, usability, and acceptability of the developed instructional model before proceeding to effectiveness testing. The small-group trial involved one fourth-grade teacher and 12 students. This initial implementation aimed to identify potential usability issues, clarity of instructions, navigation flow, and students' ability to interact with the AI-assisted features embedded in the system. Data were obtained through user response questionnaires, interface clarity feedback, technical functionality observations, and documentation of initial engagement patterns. The findings indicated that students were able to operate the system independently after brief guidance, while the teacher reported that AI-assisted feedback facilitated more efficient monitoring of student responses. Based on this limited trial, minor revisions were implemented, particularly related to navigation clarity and visual enhancement, before the model was tested in a broader classroom context.

Following refinement, the field trial was conducted in two elementary schools involving four teachers and 60 students. At this stage, structured response questionnaires were administered to assess the practicality of the product in authentic classroom settings. The questionnaire evaluated ease of use, interface attractiveness, content relevance, and perceived instructional benefits using a five-point Likert scale. The results of the practicality analysis are presented in Table 7.

Table 7. Practicality Evaluation Based on Teachers' and Students' Responses

Indicator	Teachers (M ± SD)	Category	Students (M ± SD)	Category
Ease of Use	4.50 ± 0.25	Very Practical	4.35 ± 0.30	Very Practical
Attractive Appearance	4.40 ± 0.28	Very Practical	4.42 ± 0.35	Very Practical
Relevant Content	4.55 ± 0.20	Very Practical	4.38 ± 0.32	Very Practical
Benefits for Learning	4.60 ± 0.18	Very Practical	4.45 ± 0.28	Very Practical
Total Average	**4.51 ± 0.23**	**Very Practical**	**4.40 ± 0.31**	**Very Good**

The results demonstrate that both teachers and students perceived the developed instructional model as highly practical for classroom use. Teachers emphasized the usefulness of real-time learning analytics in monitoring student progress and facilitating instructional decision-making, while students reported that the interactive elements enhanced engagement and supported their understanding of the learning material. Overall, these findings indicate that the product fulfilled the practicality criteria required for implementation before effectiveness testing.

Effectiveness of the Model

The effectiveness of the developed AI-supported deep learning model was evaluated using a one-group pretest–posttest design during the field trial stage involving 60 fourth-grade students. The analysis focused on students' meaningful literacy competencies across three dimensions: conceptual understanding, critical thinking, and knowledge application. Descriptive statistics of students' performance before and after the intervention are presented in Table 8.

Table 8. Meaningful Learning Outcomes Before and After Intervention

Learning Dimension	Pre-test (M ± SD)	Post-test (M ± SD)	Improvement (%)
Conceptual Understanding	68.45 ± 8.12	82.30 ± 7.15	20.2%
Critical Thinking	65.70 ± 9.05	80.25 ± 8.22	22.1%
Knowledge Application	63.90 ± 7.88	78.85 ± 8.15	23.4%
Total Score	**66.68 ± 8.35**	**80.47 ± 7.84**	**20.7%**

The descriptive findings indicate consistent improvement across all literacy dimensions following the intervention. The total mean score increased from 66.68 (SD = 8.35) to 80.47 (SD = 7.84), reflecting a 20.7% improvement. To determine the magnitude of improvement, normalized gain (N-Gain) scores were calculated using the formula specified in the Method section. The average normalized gain (N-Gain) for the total literacy score was 0.41, which falls into the medium category. According to the established criteria ($0.30 \leq g < 0.70$), this value falls into the medium category, indicating a moderate yet meaningful level of learning improvement. This suggests that the AI-supported deep learning model demonstrated substantial practical effectiveness in enhancing students' conceptual understanding, critical thinking, and knowledge application.

Student character development was assessed using a structured observation rubric across three indicators: cooperation, responsibility, and empathy, measured on a four-point scale (1 = very low; 4 = very high). The results are presented in Table 9.

Table 9. Students' Character Development

Character Indicator	Before Intervention (M ± SD)	After Intervention (M ± SD)	Improvement (%)
Cooperation	2.75 ± 0.65	3.90 ± 0.58	41.8%
Responsibility	2.60 ± 0.72	3.85 ± 0.61	48.0%
Empathy	2.40 ± 0.70	3.75 ± 0.55	56.3%
Total Score	**2.58 ± 0.69**	**3.83 ± 0.58**	**48.5%**

The findings show substantial improvement across all character indicators after the intervention. The total character score increased from 2.58 to 3.83, representing a 48.5% improvement. The largest increase was observed in empathy (56.3%), followed by responsibility (48.0%) and cooperation (41.8%). Based on the mean score interpretation criteria described in the Method section (3.26–4.00 = Very Good), the post-intervention character scores fall within the Very Good category. This indicates that the AI-supported deep learning model effectively fostered positive character development in terms of cooperation, responsibility, and empathy. Learning analytics data recorded student engagement over four weeks, as presented in Table 10.

Table 10. Student Engagement Based on Learning Analytics

Indicator	Week 1	Week 2	Week 3	Week 4	Improvement (%)
Interaction Frequency	320	410	480	550	71.8%
Learning Duration (minutes)	2,350	2,800	3,200	3,750	59.5%
Idea Contribution	150	190	230	300	100.0%

The data indicate a consistent upward trend in student engagement throughout the four-week implementation period. Idea contribution demonstrated the highest improvement (100%), suggesting increased active participation in AI-supported learning activities. Overall, the results demonstrate that the developed AI-supported deep learning model effectively enhanced students' meaningful literacy competencies, promoted character development, and increased learning engagement. The medium N-Gain value (g = 0.41) combined with substantial improvements in affective and engagement indicators indicates that the product meets the effectiveness criteria established in the research design and is suitable for implementation in elementary school contexts.

DISCUSSIONS

This study aimed to develop and evaluate an AI-supported deep learning model in Bahasa Indonesia instruction by examining three core criteria within Research and Development research: validity, practicality, and effectiveness. The findings demonstrate that the developed model met all three criteria empirically and theoretically, while also contributing to the integration of AI-driven personalization, meaningful literacy, and structured character education at the elementary level.

Validity of the Developed Model

The expert validation results (M = 4.28) indicate that the product falls into the "very valid" category, confirming its feasibility in terms of content relevance, instructional design coherence, linguistic clarity, and technical functionality. Within R&D paradigms, validity does not merely reflect surface-level usability but represents conceptual alignment between theoretical foundations and instructional implementation. The model developed in this study integrates deep learning stages (engage–explore–explain–elaborate–evaluate) with AI-based feedback and learning analytics. This structural coherence strengthens its pedagogical validity.

The iterative refinement process following expert recommendations (e.g., improved navigation, multimedia support, voice interaction) reflects design-based research principles emphasizing formative evaluation and contextual adaptation (Hoadley & Campos, 2022). Similar studies have reported that AI-based instructional systems require continuous calibration to ensure age-appropriate interface design and pedagogical compatibility (Holmes & Tuomi, 2022; Kim & Kwon, 2023). The present study extends these findings by demonstrating that validation procedures embedded within an R&D cycle can systematically enhance instructional robustness before large-scale implementation. Compared with prior research focusing primarily on AI functionality or technical performance (Arriola-Mendoza & Valerio-Ureña, 2024), this study emphasizes pedagogical congruence. The validation process ensured that AI was not merely an auxiliary technological tool but functioned as an instructional scaffold aligned with meaningful learning theory (Mayer, 2024). Thus, the product's validity lies in its integration of cognitive theory, character education principles, and AI-supported feedback mechanisms within a single coherent framework.

Practicality in Authentic Classroom Contexts

Practicality was confirmed through teacher and student responses in both small-group and field trials. The overall practicality scores (Teachers M = 4.51; Students M = 4.40) indicate that the model was highly practical and well-accepted in real classroom settings. Teachers highlighted the utility of real-time learning analytics for monitoring conceptual development, while students emphasized engagement derived from gamification and interactive feedback. These findings align with recent international studies indicating that AI-enhanced dashboards improve instructional decision-making efficiency and formative assessment accuracy (Fan et al., 2025; Wang, 2024). In Indonesian contexts, however, AI integration often remains at the level of content generation rather than pedagogical analytics (Efriyanti et al., 2024). The present study differs by embedding AI analytics within classroom workflow, enabling reflective teacher action and character reinforcement.

The incorporation of gamification elements within the system served to enhance its perceived practicality. Research by Carcelén-Fraile (2025) and Santos-Meneses & Drugova (2023) demonstrates that gamified digital environments enhance motivation and engagement, particularly among elementary students. Research in the field of elementary education has yielded analogous findings, indicating that the integration of interactive digital learning media has been demonstrated to have a substantial impact on enhancing motivation levels and optimising learning outcomes within classroom environments (Yonanda et al., 2024). The increasing interaction frequency and idea contribution observed in this study substantiate these theoretical claims. Notably, idea contribution increased by 100% over four weeks, suggesting that students moved from passive task completion toward active participation and meaning construction. Thus, the practicality of this model lies not only in technical usability but also in pedagogical operability. The system was manageable within existing classroom routines, compatible with the Kurikulum Merdeka emphasis on student-centered learning, and supportive of formative assessment practices.

Effectiveness in Enhancing Meaningful Learning

The effectiveness analysis demonstrated statistically significant improvement across all literacy dimensions, with an overall N-Gain of 0.41 (medium category). While categorized as moderate, this gain represents meaningful cognitive advancement within a relatively short four-week intervention. Importantly, improvement was strongest in knowledge application (23.4%), indicating that students were increasingly able to contextualize and transfer learning beyond textual recall. These findings are consistent with deep learning theory, which emphasizes application, metacognition, and conceptual integration over rote memorization (Gao, 2025; Mayer, 2024). AI-supported scaffolding likely contributed to this outcome by providing immediate corrective feedback and prompting reflective thinking. Similar effects have been observed in AI-assisted literacy instruction studies (Kosmas et al., 2025; Zhou & Peng, 2025), though many of those studies focus primarily on comprehension scores rather than integrated conceptual depth. Compared to research reporting purely motivational improvements (Boulhrir & Hamash, 2025), this study provides empirical evidence of measurable cognitive gains alongside engagement growth. The convergence between post-test improvement and rising engagement analytics strengthens internal consistency within the dataset, suggesting that increased interaction translated into deeper processing rather than superficial digital activity.

Effectiveness in Character Development

One of the most distinctive contributions of this study is its integration of character education into AI-supported literacy learning. Post-intervention character scores reached the "very good" category (M = 3.83), with empathy showing the highest improvement (56.3%). This suggests that structured reflective prompts and collaborative digital tasks may foster socio-emotional growth alongside academic skills. Previous studies have noted that digital environments can support moral reasoning and social-emotional learning when intentionally designed (Manganello & Baldacci, 2024; Zainuddin et al., 2025). However, empirical evidence combining AI analytics with structured character observation remains limited. The present findings extend this literature by demonstrating that AI-based personalization does not inherently undermine humanistic education; rather, when pedagogically framed, it can strengthen responsibility, cooperation, and empathy. Learning analytics trends further corroborate this interpretation. Increasing interaction frequency and idea contribution indicate greater student agency, which is closely linked to responsibility development. Getenet et al., (2024) emphasize that adaptive digital systems can promote participatory learning cultures when designed to encourage reflection rather than automation. The current model operationalizes this principle within Bahasa Indonesia instruction.

Theoretical and Practical Contributions

The primary theoretical contribution of this study lies in proposing an integrated triadic model combining AI-driven personalization, deep learning pedagogy, and character education within elementary literacy instruction. In contradistinction to preceding studies (Kain et al., 2024), the present research synthesises technological, cognitive and affective outcomes within a coherent teaching model, thereby providing teachers with a structured framework for embedding AI in curriculum demands. The integration of Google Forms, AI feedback, and analytics dashboards ensures scalability without requiring sophisticated infrastructure. This is particularly relevant for Indonesian elementary schools transitioning toward digital transformation under national education reforms.

CONCLUSION

This study demonstrates the distinct contribution of integrating artificial intelligence within a deep learning-based instructional framework in Bahasa Indonesia learning at the elementary school level. This approach has been shown to simultaneously enhance cognitive understanding, learning engagement, and character development. The present study is distinctive in its utilisation of a holistic AI-supported model within the context of teaching Bahasa Indonesia, and in its consideration of the cognitive, affective and social domains. The implementation enabled personalized literacy learning pathways, promoted metacognitive reflection during text interpretation, and supported students in applying linguistic and conceptual knowledge to authentic communication contexts. These findings enrich the theoretical discourse on AI integration in primary language education by providing empirical evidence that meaningful literacy learning and character formation can be cultivated together within an AI-supported Bahasa Indonesia classroom environment. Despite its substantial contributions, this study recognizes several limitations. The sample size was relatively limited and drawn from specific elementary school settings, potentially influencing generalizability across broader educational contexts. In addition, the duration of the intervention was moderately short, making it difficult to evaluate long-term literacy growth and sustained character development within Bahasa Indonesia learning. Future research should involve a wider range of participants from different regions and school types. It should also extend the duration of implementation and use more advanced analytical tools to capture deeper patterns of student literacy, reflective thinking and collaborative interaction in AI-supported language learning environments. The findings highlight the importance of integrating AI-enhanced deep learning strategies into Bahasa Indonesia instruction by encouraging teachers to adopt AI-supported literacy frameworks that promote adaptive feedback, reflective learning, contextual writing, and character-oriented discussions. In addition, policymakers and schools are encouraged to strengthen AI literacy integration, provide professional development for teachers, and establish institutional support systems to ensure responsible, effective, and sustainable AI implementation in language learning.

REFERENCES

Arriola-Mendoza, J., & Valerio-Ureña, G. (2024). Essential elements for implementing ai tools in elementary school: A systematic literature review. *Education Sciences*, *14*(12). https://doi.org/10.3390/educsci14121292

Borna, M. R., Saadat, H., Hojjati, A. T., & Akbari, E. (2024). Analyzing click data with AI: Implications For student performance prediction and learning assessment. *Frontiers in Education*, *9*(December). https://doi.org/10.3389/feduc.2024.1421479

Boulhrir, T., & Hamash, M. (2025). Unpacking artificial intelligence in elementary education: A comprehensive thematic analysis systematic review. *Computers and Education: Artificial Intelligence, 9*(June), 100442. https://doi.org/10.1016/j.caeai.2025.100442

Braun, V., & Clarke, V. (2024). Thematic analysis: A practical guide. *European Journal of Psychotherapy & Counselling*, April, 1–4. https://doi.org/10.1080/13642537.2024.2391666

Carcelén-Fraile, M. del C. (2025). Active gamification in the emotional well-being and social skills of primary education students. *Education Sciences, 15*(2). https://doi.org/10.3390/educsci15020212

Draucker, C. B., Rawl, S. M., Vode, E., & Carter-Harris, L. (2020). Integration through connecting in explanatory sequential mixed method studies. *Western Journal of Nursing Research*, *42*(12), 1137–1147. https://doi.org/10.1177/0193945920914647

Dubinsky, J. M., & Hamid, A. A. (2024). The Neuroscience of Active Learning and Direct Instruction. *Neuroscience and Biobehavioral Reviews*, *163*(May), 105737. https://doi.org/10.1016/j.neubiorev.2024.105737

Efriyanti, L., Derta, S., & Firdaus Annas. (2024). AI socialization in creating teaching materials for elementary school teachers in Bukittinggi City West Sumatra. *Salus Publica: Journal of Community Service*, *2*(1), 121–126. https://doi.org/10.58905/saluspublica.v2i1.329

Fan, G., Liu, D., Zhang, R., & Pan, L. (2025). The impact of AI-assisted pair programming on student motivation, programming anxiety, collaborative learning, and programming performance: A comparative study with traditional pair programming and individual approaches. *International Journal of STEM Education*, *12*(1). https://doi.org/10.1186/s40594-025-00537-3

Gao, Y. (2025). Deep learning-based strategies for evaluating and enhancing university teaching quality. *Computers and Education: Artificial Intelligence*, *8*(September 2024), 100362. https://doi.org/10.1016/j.caeai.2025.100362

Getenet, S., Cantle, R., Redmond, P., & Albion, P. (2024). Students' digital technology attitude, literacy and self-efficacy and their effect on online learning engagement. *International Journal of Educational Technology in Higher Education*, *21*(1). https://doi.org/10.1186/s41239-023-00437-y

González-Erena, P. V., Fernández-Guinea, S., & Kourtesis, P. (2025). Cognitive assessment and training in extended reality: Multimodal systems, clinical utility, and current challenges. *Encyclopedia, 5*(1), 8. https://doi.org/10.3390/encyclopedia5010008

Hoadley, C., & Campos, F. C. (2022). Design-based research: What it is and why it matters to studying online learning. *Educational Psychologist*, *57*(3), 207–220. https://doi.org/10.1080/00461520.2022.2079128

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education, 57*(4), 542–570. https://doi.org/10.1111/ejed.12533

Kain, C., Koschmieder, C., Matischek-Jauk, M., & Bergner, S. (2024). Mapping the Landscape: A Scoping Review of 21st Century Skills Literature in Secondary Education. *Teaching and Teacher Education, 151*(July), 104739. https://doi.org/10.1016/j.tate.2024.104739

Kemendikdasmen. (2025). *Pembelajaran Mendalam.* Pusat Kurikulum dan Pembelajaran.

Kim, K., & Kwon, K. (2023). Exploring the AI competencies of elementary school teachers in South Korea. *Computers and Education: Artificial Intelligence*, *4*(September 2022), 100137. https://doi.org/10.1016/j.caeai.2023.100137

Kosmas, P., Nisiforou, E. A., Kounnapi, E., Sophocleous, S., & Theophanous, G. (2025). Integrating artificial intelligence in literacy lessons for elementary classrooms: a co-design approach. *Educational Technology Research and Development*, 0123456789. https://doi.org/10.1007/s11423-025-10492-z

Manganello, F., & Baldacci, M. (2024). Digital stories and inclusive cultures at school: A research study in an Italian primary multicultural classroom. *Education Sciences*, *14*(10). https://doi.org/10.3390/educsci14101108

Mayer, R. E. (2024). The past, present, and future of the cognitive theory of multimedia learning. *Educational Psychology Review*, *36*(1), 1–25. https://doi.org/10.1007/s10648-023-09842-1

Moore, K., Mowers, R., Harbur, M. L., & Merrick, L. (2021). Published on plant breeding e-learning in africa the analysis of variance (ANOVA). *Plant Breeding E-Learning in Africa.* https://pbea.agron.iastate.edu

Moundridou, M., Matzakos, N., & Doukakis, S. (2024). Generative AI tools as educators' assistants: designing and implementing inquiry-based lesson plans. *Computers and Education: Artificial Intelligence*, *7*(July), 100277. https://doi.org/10.1016/j.caeai.2024.100277

Rojas, M. P., & Chiappe, A. (2024). Artificial intelligence and digital ecosystems in education: A Review. *Technology, Knowledge and Learning*, *29*(4), 2153–2170. https://doi.org/10.1007/s10758-024-09732-7

Sajja, R., Sermet, Y., Cikmaz, M., Cwiertny, D., & Demir, I. (2024). Artificial intelligence-enabled intelligent assistant for personalized and adaptive learning in higher education. *Information (Switzerland), 15*(10), 1–23. https://doi.org/10.3390/info15100596

Santos-Meneses, L. F., & Drugova, E. A. (2023). Trends in critical thinking instruction in 21st-century research and practice: Upgrading instruction in digital environments. *Thinking Skills and Creativity*, 49, 101383. https://doi.org/10.1016/j.tsc.2023.101383

Simard, A., F.-Dufour, I., Malchelosse-Fournier, R., Perreault, J., & Hudon, C. (2025). Compatibility, integration, and epistemology: Contemporary issues from a mixed methods research experiment. *Methods in Psychology*, *12*(March). https://doi.org/10.1016/j.metip.2025.100187

Soicher, R. N., Baker, A. R., & Thomas, R. C. (2024). A mixed-methods research design to advance inclusive and equitable teaching. *Innovative Higher Education*, *49*(6), 1105–1125. https://doi.org/10.1007/s10755-024-09741-5

Wang, C. (2024). Exploring students' generative AI-assisted writing processes: perceptions and experiences from native and nonnative English speakers. *Technology, Knowledge and Learning,* 0123456789. https://doi.org/10.1007/s10758-024-09744-3

Yonanda, D. A., Islahuddin, I., Ramadhani, F. A., Febriyanto, B., Saputra, D. S., Yuliati, Y., & Nurhidayat, E. (2024). Improving Motivation and Learning Outcomes of Elementary School Students with Multimedia-Based Interactive Media. *Profesi Pendidikan Dasar*, *11*(3), 197–210. https://doi.org/10.23917/ppd.v11i3.5761

Zainuddin, Z., Nafisah, A., & Muttaqin, M. (2025). Transforming character education through the implementation of the independent curriculum in Indonesia. *International Journal of Education and Humanities, 5*(3), 560–572. https://doi.org/10.58557/(ijeh).v5i3.346

Zhou, M., & Peng, S. (2025). The usage of AI in teaching and students' creativity: The mediating role of learning engagement and the moderating role of AI literacy. *Behavioral Sciences*, *15*(5). https://doi.org/10.3390/bs15050587