

# The AKSARA Learning Model: Developing an Instructional Model Through the Differentiated-Based Deep Learning (DBDL) Approach to Enhance the Literacy Competence of Junior High School Students

St. Maifah<sup>1</sup>, Munirah<sup>2\*</sup>, Aliem Bahri<sup>3</sup>

<sup>1</sup> Universitas Muhammadiyah Makassar, Makassar, South Sulawesi, Indonesia; st.maifah@gmail.com

<sup>2</sup> Universitas Muhammadiyah Makassar, Makassar, South Sulawesi, Indonesia; munirah@unismuh.ac.id

<sup>3</sup> Universitas Muhammadiyah Makassar, Makassar, South Sulawesi, Indonesia; aliembahri@unismuh.ac.id

\*Corresponding author e-mail: munirah@unismuh.ac.id

## HOW TO CITE:

St. Maifah, Munirah, Aliem Bahri  
(2026). The AKSARA Learning  
Model: Developing an Instructional  
Model Through the Differentiated-  
Based Deep Learning (DBDL)  
Approach to Enhance the Literacy  
Competence of Junior High  
School Students.  
International Journal of Special  
Education, 41(21s), 1152-1164

## ABSTRACT

This study developed and evaluated an instructional model that integrates differentiated instruction and deep learning to strengthen the literacy competence of junior high school students. The study employed a Research and Development design using the ADDIE model. The Differentiated-Based Deep Learning (DBDL) approach served as the conceptual foundation, and the development process produced the AKSARA learning model, whose five phases are Learning Readiness Analysis, Differentiated Knowledge Construction, Deep Learning Synergy, Actualization of Differentiated Meaningful Work, and Reflection and Evaluative Action. The model was supported by a model book, a teaching module, and student worksheets. Three experts validated the products and the research instruments using Aiken's V. A limited trial was conducted in one school with two Grade IX classes ( $n = 64$ ), and an expanded trial was conducted in three junior high schools in Wajo Regency, Indonesia, with six classes ( $n = 160$ ). Practicality was assessed through implementation-fidelity observation, teacher classroom-management observation, student-behaviour observation, and teacher and student response questionnaires. Change in literacy scores was examined with a one-group pretest-posttest design, Shapiro-Wilk tests, paired-sample t-tests, normalized gain, and Cohen's  $d_z$ . The model book obtained an Aiken's V of 0.961, the teaching module 0.912, and the student worksheets 0.895; the research instruments ranged from 0.897 to 0.944, and all values met the critical value of 0.889 for three raters. Practicality increased from 90.91% in the limited trial to 92.05% in the expanded trial, both in the highly practical category, and implementation fidelity rose from 91.39% in the first meeting to 94.10% in the sixth. In the expanded trial, the mean literacy score increased from 47.563 to 84.500, a mean difference of 36.938 points, 95% CI [34.730, 39.145],  $t(159) = 33.050$ ,  $p < .001$ , with an N-Gain of 0.704; 95.00% of the students achieved moderate or high individual gains. Because the design included no comparison group, because students were

## COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.

Open access publication under the  
terms and conditions

1e Creative Commons Attribution (CC BY)

license (<http://creativecommons.org/licenses/by/4.0/>).

nested in classes and schools, and because the outcome test was developed for this study and administered in identical form on both occasions, the pre-post change should be read as evidence consistent with the model rather than as an estimate of its causal effect. The findings show that differentiation and deep learning can be combined into a single syntax that experts judge valid and that teachers can execute with high fidelity in ordinary classrooms.

**Keywords:** AKSARA Learning Model, Deep Learning, Differentiated Instruction, Literacy Competence, Research and Development

## 1. Introduction

Literacy competence is the capacity to access, understand, analyse, evaluate, and use written information to make decisions and to participate in academic and social life. In Indonesia, national and international assessments have repeatedly indicated that a substantial proportion of junior high school students remain below the minimum proficiency level in reading (OECD, 2023). This pattern is not only a measurement problem; it reflects instruction that frequently treats a heterogeneous class as if it were homogeneous, and that stops at surface comprehension rather than reaching interpretation, argumentation, and transfer.

Differentiated instruction offers a principled response to learner variability. It requires teachers to adjust content, process, and product according to students' readiness, interest, and learning profile so that all learners can reach the same competence through different routes (Tomlinson, 2017). The empirical record is positive but moderate: a meta-analysis of differentiation practices reported small to moderate cognitive effects that depended strongly on how the practices were implemented (Deunk et al., 2018). Studies of reading specifically report improved comprehension under differentiated instruction (Magableh & Abdullah, 2022) and under data-based differentiation using standardized assessment with aligned material (Karst et al., 2022), and differentiated grouping has been shown to work as an instructional structure aligned with students' backgrounds rather than only with their test scores (Domke et al., 2024).

Nevertheless, differentiation on its own does not guarantee depth. Learners can be given different texts and different tasks and still remain at a reproductive level of processing. Deep learning pedagogy addresses this by emphasising cognitive engagement, meaning making, collaboration, and connection to real problems (Fullan et al., 2018). In the Indonesian policy context this orientation has been articulated as mindful, meaningful, and joyful learning, and recent conceptual and applied work has begun to describe how it can be organised at the classroom level (Andayanie et al., 2025; Kurnianingsih et al., 2026; Siregar et al., 2025). Deep processing is consistent with the wider evidence that learning improves when students are required to generate, explain, and reorganise what they have read rather than only receive it (Fiorella, 2023).

The persistent difficulty is implementation. Teachers report that differentiation increases planning workload, complicates classroom management, and is hard to sustain without concrete guidance (Bogale et al., 2026; Novriyeni et al., 2025). Development studies in Indonesia have produced differentiated products for particular subjects (Kasmawati et al., 2025), and model-development research has established procedures for testing validity, practicality, and effectiveness (Intisari et al., 2023; Rahutami et al., 2026; Sutrisno et al., 2026). What remains scarce is a single instructional model whose syntax carries both commitments at once: differentiation as the entry condition of the lesson and deep learning as the processing requirement within it, oriented explicitly toward literacy competence and supported by a complete, validated set of classroom tools.

A needs analysis conducted with teachers and Grade IX students in Wajo Regency confirmed this gap. Student needs for literacy learning experiences reached 89.68%, for differentiated learning 91.77%, and for deep learning 93.79%, all in the very high category. Teacher needs for a learning model built on the DBDL approach reached 90.00%, with readiness to implement such instruction at 87.00%, while teachers simultaneously reported that no operational device existed to guide differentiation in Indonesian language instruction.

Two terms are used consistently throughout this article and should not be conflated. Differentiated-Based Deep Learning (DBDL) is an approach: the conceptual foundation on which the development work rests. AKSARA is the learning model: the product of that development and the principal finding of the study. DBDL is therefore never referred to here as a model, and AKSARA is never referred to as an approach.

The contribution claimed here is the integration of differentiation and deep learning into one syntax in which each phase performs both functions, rather than the sequential application of two separate strategies. The study aimed to: (1) describe the needs of teachers and students that motivate the development of a learning model based on the DBDL approach; (2) produce the model prototype and its supporting devices; (3) establish the validity of the products and instruments; (4) examine the practicality of implementation; and (5) examine change in the literacy competence of Grade IX students following implementation.

## 2. Methodology

### 2.1 Design

This Research and Development study used the ADDIE model, comprising Analysis, Design, Development, Implementation, and Evaluation (Branch, 2009). Formative evaluation and revision occurred within and between stages, while summative evaluation examined the validity, practicality, and pre-post change associated with the final product. The outcome component used a one-group pretest-posttest

design in both the limited and expanded trials. No comparison group was available, so the design supports descriptive inference about change but not causal attribution; this constraint is carried through the reporting of results and is revisited in the Limitations.

During Analysis, the researchers identified learner characteristics, curriculum requirements, and instructional constraints through document study, initial observation, teacher interviews, and needs questionnaires administered to teachers and Grade IX students in Wajo Regency. During Design, the model rationale, syntax, social system, principle of reaction, support system, instructional and nurturant effects, and the assessment instruments were specified. During Development, the model book, teaching module, and student worksheets (LKPD) were produced, validated by experts, and revised. Implementation proceeded through a limited trial followed by an expanded trial after revision. Evaluation synthesised expert judgement, observation, user responses, and learning outcomes (Outhwaite et al., 2020).

### 2.2 Participants and setting

The limited trial was conducted at SMP Negeri 1 Sengkang with two Grade IX classes ( $n = 64$ ). The expanded trial was conducted in three other junior high schools in Wajo Regency: SMP Negeri 5 Sengkang (2 classes,  $n = 39$ ), SMP Negeri 1 Tanasitolo (2 classes,  $n = 60$ ), and SMP Negeri 1 Pammana (2 classes,  $n = 61$ ), giving six classes and 160 students. The Indonesian language teacher of each participating class implemented the model in that class. Each trial covered six meetings of Indonesian language instruction. Schools were selected purposively because they taught the targeted Grade IX content, represented varied student readiness profiles, and agreed to implement the full six-meeting sequence.

Because the limited and expanded trials were conducted in different schools with different student intakes, their baseline levels differ substantially (pretest means of 69.740 and 47.563 respectively). The two trials are therefore reported separately throughout, are not pooled,

and should not be read as successive measurements of a single sample. The difference in baseline is a property of the samples, not of a change in the instrument, which was identical in both trials.

Permission for the study was obtained from the Wajo Regency Education and Culture Office and from each school. Teachers gave informed consent, and student participation was approved by the schools acting in loco parentis, with students informed that test results would be used for research purposes and reported anonymously.

### 2.3 The model and its supporting devices

Each lesson followed the five-phase AKSARA syntax with an opening and a closing segment (see Section 3.1). The support system comprised the model book, the teacher guide, the student book, the teaching module, the student worksheets, learning media, and the assessment instruments.

### 2.4 Instruments

Validation involved three experts, selected for complementary expertise in (a) Indonesian language instruction, literacy, and deep learning; (b) instructional-model development and differentiated learning; and (c) learning devices and educational evaluation. Two groups of objects were validated: the development products (model book, teaching module, student worksheets) and the research instruments (implementation-fidelity observation sheet, literacy competence test, teacher response questionnaire, student response questionnaire).

Data were collected using needs questionnaires and interviews, expert validation sheets, implementation-fidelity observation sheets, teacher classroom-management observation sheets, student-behaviour observation sheets, teacher and student response questionnaires, and a 30-item literacy competence test. The test represented accessing information, understanding texts, analysing and evaluating content, interpreting meaning, and using information to solve contextual problems. The same form was administered as pretest and posttest, separated by the six-meeting sequence. Content validity of the test was established

through expert judgement; internal-consistency reliability and item-level difficulty and discrimination indices were not computed, which limits the psychometric claims that can be made for the outcome measure.

Observations were conducted by the researcher together with the classroom teacher using the validated observation sheets. Observers were not blind to the purpose of the study, and interrater agreement was not quantified; both conditions are treated as limitations rather than as neutral procedural details.

### 2.5 Analysis

Expert ratings were analysed with Aiken's V (Aiken, 1985) on a four-point rating scale with three raters. Rather than applying descriptive bands alone, this study adopted the critical value of V for this rater and category configuration, so that an item or aspect was judged valid when  $V \geq 0.889$ ; critical values for V were interpreted following Penfield and Giacobbi (2004). Coefficients are reported for each object separately and are not averaged across objects, because an average of V across heterogeneous instruments has no defensible interpretation.

Practicality was computed as the percentage of the obtained score relative to the maximum score, with 81%–100% classified as highly practical. Pretest and posttest scores were summarised descriptively and by category distribution. Shapiro-Wilk tests were applied to the pretest, posttest, and difference scores; because the difference scores were normally distributed in both trials, paired-sample t-tests were used. Normalized gain was calculated as  $(\text{posttest} - \text{pretest}) / (\text{maximum} - \text{pretest})$  and interpreted with the criteria  $g \geq 0.70$  (high),  $0.30 \leq g < 0.70$  (moderate), and  $g < 0.30$  (low) (Hake, 1998). Effect magnitude was estimated with Cohen's *d*. Statistical significance was set at .05.

Students were nested within classes and schools, but the analyses treat students as independent observations because outcome data were recorded at student level only and were not disaggregated by class. Where intra-class correlation is non-zero, this procedure understates the standard errors and therefore

overstates the test statistics; the significance values reported below should be read with that qualification, and multilevel analysis is recommended for replication.

### 3. Results

#### 3.1 Needs analysis and model prototype

The needs analysis produced consistently very high scores for both groups of users. Students reported the highest need for deep learning experiences, while teachers reported the highest need for a model that could operationalise differentiation in Indonesian language instruction (Table 1).

**Table 1. Needs analysis of students and teachers**

| Respondent | Indicator                                                     | Percentage | Category  |
|------------|---------------------------------------------------------------|------------|-----------|
| Students   | Need for literacy learning experience in Indonesian language  | 89.68      | Very high |
| Students   | Need for differentiated learning (content, process, product)  | 91.77      | Very high |
| Students   | Need for deep learning (mindful, meaningful, joyful)          | 93.79      | Very high |
| Teachers   | Orientation to the Merdeka Curriculum and literacy competence | 88.67      | Very high |
| Teachers   | Need for a learning model based on the DBDL approach          | 90.00      | Very high |
| Teachers   | Need for literacy instruction                                 | 89.33      | Very high |
| Teachers   | Readiness to implement DBDL-based instruction                 | 87.00      | Very high |

Note. Percentages are of the maximum attainable questionnaire score. Values of 81%–100% are classified as very high.

These findings were translated into design decisions. Because learner variability was the recurring theme in both data sets, diagnostic identification of readiness, interest, and learning profile was placed at the beginning of the lesson rather than at the point of remediation. Because students requested meaning rather than coverage, the middle of the lesson was reserved for dialogue, analysis, and argumentation. Because teachers reported the absence of operational guidance, each phase was accompanied by

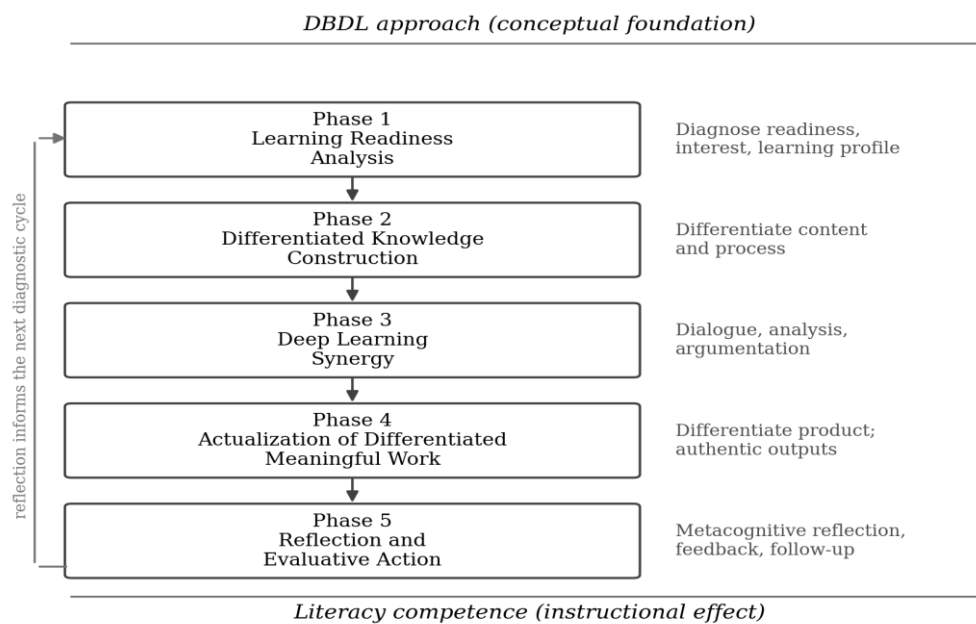
concrete devices: a model book, a teaching module, and student worksheets.

The development process produced the AKSARA learning model (AKtualisasi Sintaks pembelajARan berdiferensiAsi berbasis Deep Learning), whose five phases are specified in Table 2 and whose internal logic is shown in Figure 1. DBDL is the conceptual approach that guided the development; AKSARA is the resulting model and the principal finding of the study.

**Table 2. Syntax of the AKSARA learning model**

| Phase | Name                                            | Teacher action                                                                       | Intended literacy function                                             |
|-------|-------------------------------------------------|--------------------------------------------------------------------------------------|------------------------------------------------------------------------|
| 1     | Learning Readiness Analysis                     | Diagnostic assessment of readiness, interest, and learning profile                   | Establishes each student's entry point in reading and reasoning        |
| 2     | Differentiated Knowledge Construction           | Differentiation of content and process through varied texts, media, and routes       | Provides equitable access to text at appropriate complexity            |
| 3     | Deep Learning Synergy                           | Facilitation of dialogue, analysis, argumentation, and collaboration around the text | Moves comprehension from retrieval to interpretation and evaluation    |
| 4     | Actualization of Differentiated Meaningful Work | Supervision of product differentiation through authentic student-selected outputs    | Requires students to transform understanding into communicated meaning |
| 5     | Reflection and Evaluative Action                | Guidance of metacognitive reflection, feedback, and follow-up planning               | Consolidates literacy strategies and self-regulation                   |

Note. The table states the design specification of the model. The intended literacy functions are design claims, not empirical results.



**Figure 1. Phase structure of the AKSARA learning model within the DBDL approach**

The model was completed by its social system (student-centred, dialogic, collaborative), principle of reaction (teacher as diagnostician and facilitator responding differently in each phase), support system (model book, teacher guide, student book, teaching module, worksheets, media, and assessment instruments), instructional effects (literacy competence, higher-order thinking, authentic work), and nurturant effects (self-directed learning, collaboration, creativity,

reflective attitude, appreciation of diversity, intrinsic motivation).

### 3.2 Validity

All development products and research instruments met the critical Aiken's V value of 0.889 (Table 3). The model book obtained the highest coefficient, with full agreement among the three experts on the benefit and language aspects.

**Table 3. Expert validation coefficients by object**

| Object of validation                      | Aiken's V | Decision             |
|-------------------------------------------|-----------|----------------------|
| Model book                                | 0.961     | Meets critical value |
| Teaching module                           | 0.912     | Meets critical value |
| Student worksheets (LKPD)                 | 0.895     | Meets critical value |
| Implementation-fidelity observation sheet | 0.944     | Meets critical value |
| Literacy competence test                  | 0.910     | Meets critical value |
| Teacher response questionnaire            | 0.915     | Meets critical value |
| Student response questionnaire            | 0.897     | Meets critical value |

Note. Three raters, four-point scale; critical value  $V \geq 0.889$  (Aiken, 1985; Penfield & Giacobbi, 2004). Coefficients are reported per object and are not averaged across objects.

Validation was also used diagnostically at the item level. Three indicators in the teaching module fell below the critical value, principally the specification of learning objectives from the curriculum achievement statements ( $V = 0.778$ ) and the alignment of methods with competence indicators ( $V = 0.778$ ). Four indicators in the student worksheets fell below the critical value, the lowest being visual layout and illustration ( $V = 0.667$ ). One construction item in each response questionnaire obtained  $V = 0.778$  because it measured more than one idea. Each of these items was revised: objectives were rewritten with operational verbs and mapped to literacy indicators, activities were re-sequenced so that each method served a stated indicator, worksheet layout and illustrations were redesigned, collaborative and reflective tasks were added, and double-barrelled questionnaire items were split into single-idea statements. No revision altered the five-phase syntax.

### 3.3 Practicality

Practicality was evaluated from five data sources in both trials. In the limited trial,

implementation fidelity rose across the six meetings from 85.94% in the first meeting to 95.24% in the sixth, with a mean of 90.87%. Teacher classroom management averaged 92.19% and student behaviour during learning 90.33%; teacher responses reached 88.89% and student responses 92.25%. The combined practicality score was 90.91%.

In the expanded trial, implementation fidelity averaged 92.61% and again rose across meetings, from 91.39% to 94.10%. School-level means were 95.10% at SMP Negeri 5 Sengkang, 91.53% at SMP Negeri 1 Pammana, and 91.18% at SMP Negeri 1 Tanasitolo. Teacher classroom management averaged 93.08%, with the highest score in Phase 3, Deep Learning Synergy (95.14%), and student behaviour during learning averaged 92.41%. The combined practicality score was 92.05% (Table 4). Four of the five indicators were higher in the expanded trial than in the limited trial, and the fifth was 0.55 points lower.

**Table 4. Practicality indicators in the limited and expanded trials**

| Indicator                    | Limited trial (%) | Expanded trial (%) | Trend     | Category         |
|------------------------------|-------------------|--------------------|-----------|------------------|
| Implementation fidelity      | 90.87             | 92.61              | Increased | Highly practical |
| Teacher response             | 88.89             | 90.44              | Increased | Highly practical |
| Student response             | 92.25             | 91.70              | Stable    | Highly practical |
| Teacher classroom management | 92.19             | 93.08              | Increased | Highly practical |
| Observed student behaviour   | 90.33             | 92.41              | Increased | Highly practical |
| Overall practicality         | 90.91             | 92.05              | Increased | Highly practical |

Note. Limited trial: one school, two classes,  $n = 64$ . Expanded trial: three schools, six classes,  $n = 160$ . Percentages are of the maximum attainable score; 81%–100% is classified as highly practical. Observers were not blind to the intervention and interrater agreement was not quantified.

### 3.4 Change in literacy competence

In the limited trial ( $n = 64$ ), the mean literacy score increased from 69.740 ( $SD = 7.728$ ) to 92.344 ( $SD = 4.471$ ). The difference scores were normally distributed (Shapiro-Wilk  $p = .342$ ) and the paired comparison was significant, mean difference = 22.604,  $t(63) = 21.722$ ,  $p < .001$ ,

with Cohen's  $d_z = 2.72$ . The mean N-Gain was 0.739; 44 students (68.75%) achieved high gain, 17 (26.56%) moderate gain, and 3 (4.69%) low gain. Scores in this trial were compressed at the upper end of the scale: the posttest minimum was 80.00 and 61 of 64 students (95.31%) fell in the highest band, so both the normalized gain and the effect size for this trial are affected by a ceiling.

In the expanded trial ( $n = 160$ ), the mean increased from 47.563 ( $SD = 11.944$ ) to 84.500 ( $SD = 7.745$ ). The difference scores met the normality criterion (Shapiro-Wilk  $p = .214$ ), and the paired comparison was significant, mean

difference = 36.938 ( $SD = 14.137$ ), 95% CI [34.730, 39.145],  $t(159) = 33.050$ ,  $p < .001$ , Cohen's  $d_z = 2.61$ . The standard deviation fell from 11.944 to 7.745 (Table 5).

**Table 5. Literacy scores before and after implementation**

| Statistic       | Limited trial ( $n = 64$ ) | Expanded trial ( $n = 160$ ) |
|-----------------|----------------------------|------------------------------|
| Pretest M (SD)  | 69.740 (7.728)             | 47.563 (11.944)              |
| Posttest M (SD) | 92.344 (4.471)             | 84.500 (7.745)               |
| Mean difference | 22.604                     | 36.938                       |
| t (df)          | 21.722 (63)                | 33.050 (159)                 |
| p               | < .001                     | < .001                       |
| Cohen's $d_z$   | 2.72                       | 2.61                         |
| N-Gain          | 0.739 (high)               | 0.704 (high)                 |

Note. The two trials were conducted in different schools and are not pooled. Students were nested in classes and schools; clustering is not modelled, so  $p$  values are anti-conservative. Effect sizes are within-subject and are inflated by ceiling effects, particularly in the limited trial.

The distribution of normalized gain in the expanded trial showed that 97 students (60.63%) achieved high gain, 55 (34.37%) moderate gain, and 8 (5.00%) low gain; thus 95.00% of participants improved at a moderate or high level. The category distribution of scores changed in the same direction: before instruction, 1 student (0.63%) was in the very low band, 52 (32.50%) in

the low band, 91 (56.88%) in the moderate band, 16 (10.00%) in the high band, and none in the very high band; after instruction, 114 students (71.25%) reached the very high band and 45 (28.13%) the high band, leaving 1 student (0.63%) in the moderate band and none below it (Table 6).

**Table 6. Distribution of literacy scores before and after instruction (expanded trial,  $n = 160$ )**

| Score interval | Category  | Pretest f (%) | Posttest f (%) |
|----------------|-----------|---------------|----------------|
| 81–100         | Very high | 0 (0.00)      | 114 (71.25)    |
| 61–80          | High      | 16 (10.00)    | 45 (28.13)     |
| 41–60          | Moderate  | 91 (56.88)    | 1 (0.63)       |
| 21–40          | Low       | 52 (32.50)    | 0 (0.00)       |
| 0–20           | Very low  | 1 (0.63)      | 0 (0.00)       |

Note. Scores are percentages of the maximum test score. Bands are those used in the source study.

#### 4. Discussion

The expert judgements indicate alignment among the model rationale, the syntax, the teaching devices, and the assessment instruments, which is the condition under which a model can be said to be internally coherent rather than merely well described (Intisari et al., 2023; Sutrisno et al., 2026).

The validity evidence is more informative at the item level than at the aggregate level. The aspects that fell below the critical value were not conceptual; they concerned the translation of the model into usable material, namely objective specification, method-indicator alignment, visual layout, and single-idea questionnaire wording. This distribution suggests that the theoretical integration of differentiation and deep learning

was accepted by all three experts, while the operational surface required refinement. Treating the critical value of Aiken's *V* as a revision trigger, rather than treating a high mean as sufficient, made this diagnosis possible (Aiken, 1985; Penfield & Giacobbi, 2004).

Practicality can be explained by the structure of the syntax itself. Each phase specifies what the teacher diagnoses, provides, facilitates, supervises, and evaluates, so the differentiation decisions that teachers usually report as burdensome are distributed across the lesson instead of concentrated in preparation (Bogale et al., 2026; Novriyeni et al., 2025). The upward trend in implementation fidelity within both trials, from 85.94% to 95.24% and from 91.39% to 94.10%, is consistent with a learnable routine, since teachers improved with repetition without additional training between meetings. The rise in fidelity within each trial, combined with the rise in teacher classroom-management scores between trials, indicates that the syntax became easier to execute with repetition rather than more burdensome. The highest classroom-management score in the expanded trial occurred in Phase 3, the deep learning phase, which is the phase most often reported as difficult to sustain in ordinary classrooms.

The pre-post differences and the high normalized gains are consistent with the intended operation of the model, and three design mechanisms are plausible candidates. First, Phase 1 required an explicit diagnosis before any text was assigned, so that differentiation determined the starting point rather than compensating after failure; this is the condition under which data-based differentiation has previously shown comprehension gains (Karst et al., 2022; Magableh & Abdullah, 2022). Second, Phases 3 and 4 required students to argue about and then reproduce meaning in a product of their choice, that is, to generate rather than to receive, which is associated with durable comprehension (Fiorella, 2023; Fullan et al., 2018). Third, Phase

5 made reflection a scheduled component rather than an optional closing, supporting the metacognitive control that literacy tasks demand.

These mechanisms are, however, hypotheses consistent with the data rather than findings established by it. The observed effect sizes ( $d_z = 2.61$  and  $2.72$ ) are far larger than those typically reported for differentiation interventions, where meta-analytic estimates are small to moderate (Deunk et al., 2018). Three features of the present design plausibly account for much of that gap: the absence of a comparison group, so that maturation, repeated testing, teacher effects, and novelty are not separated from the effect of the model; the use of an identical test form on both occasions after a short interval; and compression at the top of the scale, which is explicit in the limited trial, where the posttest minimum was 80.00. A study of this design can establish that students' scores rose substantially under the model; it cannot establish how much of that rise the model caused.

The reduction in standard deviation from 11.944 to 7.745 in the expanded trial is nevertheless notable, because it is not what a ceiling effect alone would produce in a group whose posttest mean was 84.500 with 28.13% of students still below the highest band. The proportion of students in the two lowest bands fell from 33.13% to zero. This is the outcome that differentiation is intended to produce, and it is at least consistent with the claim that the model reaches the weaker part of the distribution rather than only the stronger.

Table 7 positions the model against the instructional models that were reviewed during its design. The comparison is a design analysis rather than an empirical benchmark: no head-to-head trial against these models was conducted, and the ratings express the extent to which each model's published syntax makes the three commitments explicit.

**Table 7. Design positioning of AKSARA relative to reviewed instructional models**

| Model                         | Differentiation in the syntax | Deep processing in the syntax | Explicit literacy orientation |
|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
| <b>Project-Based Learning</b> | Partial                       | Explicit                      | Partial                       |
| <b>Problem-Based Learning</b> | Partial                       | Partial                       | Explicit                      |
| <b>Collaborative Inquiry</b>  | Partial                       | Explicit                      | Partial                       |
| <b>AKSARA (this study)</b>    | Explicit (Phases 1, 2, 4)     | Explicit (Phases 3, 5)        | Explicit (all phases)         |

Note. Design analysis based on the published syntax of each model; not an empirical comparison. "Partial" indicates that the commitment is possible but not required by the syntax.

Several limitations constrain interpretation. First, the design included no concurrent control group, so causal attribution is not warranted. Second, the outcome test was developed for this study; content validity was established, but internal-consistency reliability and item statistics were not computed, and the same form was used at pretest and posttest. Third, ceiling effects are present, decisively in the limited trial. Fourth, students were nested in classes and schools, and clustering was not modelled, so the reported p values are anti-conservative. Fifth, the two trials differ substantially at baseline (69.740 versus 47.563) because they were conducted in different schools, so they are reported separately and cross-trial inference is not attempted. Sixth, outcome data were not disaggregated by school, so the relationship between implementation fidelity, which did vary by school, and learning gain could not be tested; this is the analysis that would most strengthen the causal argument and it should be planned into replications. Seventh, fidelity, classroom management, and student behaviour were rated by observers aware of the intervention, without quantified interrater agreement. Finally, the content was limited to Indonesian language instruction at Grade IX in one regency.

The practical implication is that differentiation and deep learning need not be implemented as two competing initiatives. When they are combined in a single syntax with matching devices, teachers in ordinary schools executed the model at above 90% fidelity from the first meeting onward, using no equipment beyond the printed and digital materials already available to them. That claim rests on the validity

and practicality evidence, which the design supports, rather than on the magnitude of the pre-post gain, which it does not.

## 5. Conclusion

This study developed the AKSARA learning model through the Differentiated-Based Deep Learning approach to strengthen the literacy competence of junior high school students in Wajo Regency. The model consists of five phases, Learning Readiness Analysis, Differentiated Knowledge Construction, Deep Learning Synergy, Actualization of Differentiated Meaningful Work, and Reflection and Evaluative Action, and is supported by a model book, a teaching module, and student worksheets. Development followed the ADDIE stages, so that the prototype was revised continuously in response to user needs, expert judgement, and trial findings.

Expert validation produced Aiken's V coefficients of 0.961 for the model book, 0.912 for the teaching module, and 0.895 for the student worksheets, with instrument coefficients between 0.897 and 0.944, all above the critical value of 0.889 for three raters on a four-point scale. Practicality increased from 90.91% in the limited trial to 92.05% in the expanded trial, with implementation fidelity rising within both trials and remaining above 91% in every school. In the expanded trial with 160 students from three schools, the mean literacy score increased from 47.563 to 84.500,  $t(159) = 33.050$ ,  $p < .001$ , with an N-Gain of 0.704, and score dispersion decreased. Given the absence of a comparison group, the use of a single test form on both occasions, unmodelled clustering, and ceiling

effects, this change is reported as evidence consistent with the model rather than as an estimate of its causal effect.

The contribution of this study is a syntax in which differentiation and deep learning operate simultaneously rather than sequentially: learner variability is diagnosed before content is assigned, and depth is required within the differentiated route rather than added afterwards. Priorities for further research follow directly from the limitations: a controlled or quasi-experimental

design with a comparison group; a psychometrically documented outcome measure with parallel forms and reported reliability; multilevel analysis that respects the nesting of students in classes and schools; disaggregated outcome data permitting a test of the fidelity-gain relationship; retention testing; and implementation in other subjects, grade levels, and regions, including settings where teachers adopt the model without researcher support.

## References

- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1), 131–142. <https://doi.org/10.1177/0013164485451012>
- Andayanie, L. M., Adhantoro, M. S., Purnomo, E., & Kurniaji, G. T. (2025). Implementation of deep learning in education: Towards mindful, meaningful, and joyful learning experiences. *Journal of Deep Learning*, 1(1), 47–56. <https://doi.org/10.23917/jdl.v1i1.11157>
- Bogale, M. R., Mekuria, Y. S., Abo, G. B., & Zinabu, S. A. (2026). Teachers' beliefs, practices, and challenges in implementing differentiated instruction in Ethiopia. *Discover Education*, 5(1). <https://doi.org/10.1007/s44217-026-01919-7>
- Branch, R. M. (2009). *Instructional design: The ADDIE approach*. Springer. <https://doi.org/10.1007/978-0-387-09506-6>
- Deunk, M. I., Smale-Jacobse, A. E., de Boer, H., Doolaard, S., & Bosker, R. J. (2018). Effective differentiation practices: A systematic review and meta-analysis of studies on the cognitive effects of differentiation practices in primary education. *Educational Research Review*, 24, 31–54. <https://doi.org/10.1016/j.edurev.2018.02.002>
- Domke, L. M., Kaplan, M., & Bingham, G. E. (2024). Differentiated reading groups: Bridging literacy instruction and culturally sustaining pedagogy. *The Reading Teacher*, 77(6), 800–809. <https://doi.org/10.1002/trtr.2317>
- Fiorella, L. (2023). Making sense of generative learning. *Educational Psychology Review*, 35(2). <https://doi.org/10.1007/s10648-023-09769-7>
- Fullan, M., Quinn, J., & McEachen, J. (2018). *Deep learning: Engage the world change the world*. Corwin.
- 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>
- Idris, M., Sugita, G., Anggraini, A., Ahmar, A. S., & Rahman, A. (2026). Evaluating the individual and combined effects of learning discipline, student independence, and parental attention on mathematics achievement among Indonesian Grade XI science students. *Discover Education*, 5(1), 1–26. <https://doi.org/10.1007/s44217-026-01310-6>
- Intisari, Syam, H., & Gani, H. A. (2023). Validity, practicality, and effectiveness of character learning model based on Al-Islam Kemuhammadiyaan values. *Asian Journal of Education and Social Studies*, 48(1), 37–49. <https://doi.org/10.9734/ajess/2023/v48i11037>
- Joyce, B., Weil, M., & Calhoun, E. (2015). *Models of teaching* (9th ed.). Pearson.

- Karst, K., Bonefeld, M., Dotzel, S., Fehringer, B. C. O. F., & Steinwascher, M. (2022). Data-based differentiated instruction: The impact of standardized assessment and aligned teaching material on students' reading comprehension. *Learning and Instruction*, 79, 101597. <https://doi.org/10.1016/j.learninstruc.2022.101597>
- Kasmawati, Nasution, E. S., & Nasution, F. (2025). Development of digital interactive worksheets based on a differentiated instruction model in science education to enhance scientific literacy in physics among junior high school students. *Jurnal Penelitian Pendidikan IPA*, 11(9), 462–472. <https://doi.org/10.29303/jppipa.v11i9.12519>
- Kurnianingsih, E., Prayitno, H. J., Mu'ti, A., Baidhaw, Z., & Jufriansah, A. (2026). Deep learning as an integrated pedagogical framework: Synthesizing mindful, meaningful, and joyful learning for transformative educational experiences. *Journal of Deep Learning*, 2(2), 125–143. <https://doi.org/10.23917/jdl.v2i2.16008>
- Magableh, I. S., & Abdullah, A. (2022). Differentiated instruction effectiveness on the secondary stage students' reading comprehension proficiency level in Jordan. *International Journal of Evaluation and Research in Education*, 11(1), 459–466. <https://doi.org/10.11591/ijere.v11i1.21971>
- Novriyeni, F., Roza, Y., & Saragih, S. (2025). Teachers' readiness and challenges in implementing differentiated instruction in mathematics education. *AXIOM: Jurnal Pendidikan dan Matematika*, 14(2), 141. <https://doi.org/10.30821/axiom.v14i2.20027>
- OECD. (2023). PISA 2022 results (Volume I): The state of learning and equity in education. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Outhwaite, L. A., Gulliford, A., & Pitchford, N. J. (2020). A new methodological approach for evaluating the impact of educational intervention implementation on learning outcomes. *International Journal of Research and Method in Education*, 43(3), 225–242. <https://doi.org/10.1080/1743727X.2019.1657081>
- Penfield, R. D., & Giacobbi, P. R. (2004). Applying a score confidence interval to Aiken's item content-relevance index. *Measurement in Physical Education and Exercise Science*, 8(4), 213–225. [https://doi.org/10.1207/s15327841mpee0804\\_3](https://doi.org/10.1207/s15327841mpee0804_3)
- Rahutami, Diana, L. N., Wadji, & Himawan, R. (2026). Developing an ADDIE-based TGTW learning model for editorial text writing: An evaluation of its validity, practicality, and effectiveness. *Journal of Language and Literature Studies*, 6(2), 547–560. <https://doi.org/10.36312/jolls.v6i2.4833>
- Siregar, T., Fauzan, A., Yerizon, & Syafriandi. (2025). Designing mathematics teaching through deep learning pedagogy: Toward meaningful, mindful, and joyful learning. *Journal of Deep Learning*, 1(2), 188–202. <https://doi.org/10.23917/jdl.v1i2.11969>
- Sutrisno, T., Suprpto, N., Mariana, N., & Suryanti. (2026). Validity and practicality of the Asi<sup>6</sup> model as an innovative learning model in Indonesia: Integration of contextual learning and local wisdom. *Multidisciplinary Science Journal*, 8(12), 2026762. <https://doi.org/10.31893/multiscience.2026762>
- Tomlinson, C. A. (2017). *How to differentiate instruction in academically diverse classrooms* (3rd ed.). ASCD.

## 7. Acknowledgements

We gratefully acknowledge the doctoral promoters, the expert validators, and the participating schools, teachers, and students for their guidance, support, and participation throughout the development and trial of the learning model. We also appreciate the institutional assistance that facilitated the implementation of this study.

## 8. Author Declaration

**Originality.** We declare that this manuscript is an original work derived from our dissertation research and has not been published or submitted elsewhere. Related components of the same development study, namely the needs analysis and the implementation and practicality analysis, are reported in separate manuscripts, and this overlap has been disclosed to the editor.

**Ethics.** Permission was obtained from the Wajo Regency Education and Culture Office and from the participating schools. Teachers provided informed consent, and student participation was approved by the schools acting in loco parentis. All data were recorded and reported anonymously.

**Funding.** This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Data availability.** The de-identified pretest and posttest scores, observation records, and questionnaire data that support the findings are available from the corresponding author on reasonable request.

**Conflicts of interest.** The authors declare no conflicts of interest related to this study.

**Author contributions.** All authors have reviewed and approved the final manuscript.