

Research Article

Cite this article: Subyantoro, Siroj, M. B., Putra, L. V., Kusriani, E., & Septiani, S. (2026). Adaptive Design for Critical Literacy: Validation of Readability-Based Reading Materials and HOTS Assessment in Indonesian Education (Grades 4–12). *Educational Process: International Journal*, 23, e2026073. <https://doi.org/10.22521/edupij.2026.23.73>

Received October 5, 2025


Accepted December 4, 2025

Keywords: Higher-Order Thinking Skills, Literacy, Readability

Author for correspondence:

Subyantoro

 bintoro@mail.unnes.ac.id

 Faculty of Language and Arts, Universitas Negeri Semarang, Indonesia

Adaptive Design for Critical Literacy: Validation of Readability-Based Reading Materials and HOTS Assessment in Indonesian Education (Grades 4–12)

Subyantoro , Muhammad Badrus Siroj , Lisa Virdinarti Putra , Endang Kusriani , Sisca Septiani 

Abstract

Background/purpose. This study developed and validated adaptive reading materials based on readability integrated with Higher Order Thinking Skills (HOTS) assessment for Indonesian elementary and secondary school students (grades 4–12).

Materials/methods. This study used a Research and Development (R&D) design. The materials were reconstructed based on the Flesch–Kincaid Grade Level (FKGL) linear progression for grades 4–12, and the HOTS (C3–C6) assessment was designed to be calibrated. Validation was conducted by various stakeholders (N≈3000 students, N≈250 teachers, N=12 experts). Effectiveness was tested using a quasi-experimental design (Paired Sample T-test) with a sample of N=30, supplemented by item psychometric analysis (Difficulty and Discrimination Indices). Using a Research and Development (R&D) design based on the Borg & Gall model, this study involved multi-stakeholder validation with approximately 3,000 students, 250 teachers, and 12 experts, followed by a quasi-experimental effectiveness test (N=30).

Results. The results show high design validity and successful text calibration with Flesch Reading Ease scores ranging from 85.2 (Grade 4) to 45.1 (Grade 12). High acceptance from stakeholders (average score >4.0/5.0). Psychometric analysis confirmed the difficulty level of most HOTS items (C3–C5). Effectiveness testing proved a significant improvement in critical reading competency (average increase of 12.567 points; p=0.000).

Conclusion. This integrated adaptive design model has proven to be valid, feasible, and significantly effective in optimising Critical Reading Competency. This product offers a data-driven and scalable benchmark for improving critical literacy in Indonesian primary and secondary education.



OPEN ACCESS

© The Author(s), 2025. This is an Open Access article, distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction, provided the original article is properly cited.

1. Introduction

Critical reading competence is a fundamental pillar of modern education systems, serving as a prerequisite for mastering Higher-Order Thinking Skills (HOTS) (Mardiana et al., 2025). In Primary and Secondary Education in Indonesia (Grades 4 to 12), there is a great urgency to improve students' literacy and critical thinking skills (Weng, 2023). This research is based on the need to address curricular gaps, where available teaching materials are often not systematically calibrated to support measurable cognitive progression (Maxnun et al., 2024). Improving critical reading competence requires materials that are not only thematically relevant but also scientifically designed to ensure that textual complexity (cognitive input) increases in line with students' cognitive development (Foo, 2025).

This competency encompasses not only surface-level comprehension but also the ability to critically analyze, evaluate, and synthesize information (Ardhian et al., 2020). In Indonesian education, this competency remains a priority (Belda-Medina, 2024), but it also presents an ongoing challenge. Evidence from the Program for International Student Assessment (PISA) shows that Indonesia consistently scores below the OECD average in reading literacy. In PISA 2022, Indonesia ranked 68th globally with an average reading score of 371, reflecting limited progress over the past two decades (Ulkhay, 2024). These alarming results clearly highlight the urgency of pedagogical innovation that not only improves basic comprehension but also fosters higher-order thinking skills (HOTS) (Lu et al., 2021).

Compounding this challenge, Indonesian classrooms must serve a diverse student population across the Primary and Secondary Education system, which encompasses Grades 4 through 12, where reading materials are often not aligned with students' cognitive readiness (Adam et al., 2025). Texts that are too simple fail to stimulate critical engagement, while texts that are too complex risk disengagement and surface-level learning (Li et al., 2024). Addressing this misalignment requires an approach based on measurable learning scaffolding. To achieve this, readability metrics such as Flesch Reading Ease (FRE) and Flesch-Kincaid Grade Level (FKGL) provide a quantitative means to calibrate text complexity to students' developmental stages (Li et al., 2024), a strategy that previous studies have confirmed can improve comprehension and ensure adequate cognitive challenge.

Parallel to the need to control linguistic input, HOTS assessment is essential for measuring and developing critical reading output (Puspitadani et al., 2022). Based on Bloom's revised taxonomy, HOTS tasks go beyond basic recall to challenge students in analysis (C3), evaluation (C4), synthesis (C5), and creation (C6), thereby requiring them to engage critically with the text (Pujawan et al., 2022). This indicates a significant research gap: no model systematically integrates data-based text calibration (FRE/FKGL) with rigorously verified HOTS question design (C3-C6) at all levels of basic education (grades 4-12).

This study offers novelty and originality through an integrated, double-layered innovation: first, the reconstruction of adaptive reading materials to meet readability metrics (FRE/FKGL) across nine grade levels (grades 4–12); and second, the systematic design of HOTS-based assessments (C3–C6) calibrated to these reconstructed texts. The validation of this study goes beyond content validity by including Psychometric Validation of HOTS items and Empirical Evidence of Effectiveness (Pre-Test/Post-Test) on a large population scale, which is an important contribution to the critical literacy literature. The alignment of this study with the global pedagogical shift towards deep learning approaches underscores the urgency of the research, as critical reading and HOTS-based tasks are cornerstones of deep learning.

The main objective of this study is to examine the impact of reading materials based on readability and HOTS assessment on critical reading skills among elementary and secondary school students in Indonesia (grades 4-12).

The specific research questions are: (1) How are adaptive reading materials reconstructed and calibrated to ensure measurable progress in text complexity? (2) How valid and feasible are these materials according to multi-party assessments? (3) How significant is the empirical improvement in Critical Reading Skills after the intervention?

2. Literature Review

2.1. Critical Reading Competency

Critical reading competency is a multifaceted skill essential for engaging with complex texts and social issues (Gil-Rojas et al., 2025). It is directly related to the development of critical thinking, in which learners interpret and critically evaluate information, thereby improving their competency. Students with higher critical reading skills demonstrate better analytical thinking and are better prepared to participate in the democratic process (Barton et al., 2024). Furthermore, reading competence correlates with broader learning outcomes, including academic performance and civic engagement, underscoring the importance of cultivating these skills in early education (Rahmawati et al., 2023). An effective educational framework should foster critical reading by integrating strategies that promote critical questioning, analysis, and interdisciplinary literacy, thereby preparing students for informed participation in a complex world (Ardhian et al., 2020; Stepanenko et al., 2023).

2.2. Readability in Literacy Development

Readability measures, particularly Flesch Reading Ease (FRE) and Flesch-Kincaid Grade Level (FKGL), effectively assess text complexity (Beck et al., 2025; Panigrahi & Bakshi, 2021; Randell et al., 2025). Using these readability tests can help create content that is easier to understand and more engaging (Burke et al., 2024). In instructional design, these formulas serve a transformative function: teachers can use FRE scores to select age-appropriate texts (e.g., FRE scores of 80-90 for elementary school students and 50-60 for middle school students), while curriculum developers can systematically classify reading materials by educational level. For example, a 5th-grade science text with an FRE of 75 ensures accessibility while maintaining an appropriate cognitive challenge, whereas the same content, rewritten with an FRE of 55, would be suitable for 9th-grade students. (Yüceer-Çetiner et al., 2025).

2.3. Higher Order Thinking Skills (HOTS) in Reading

Higher-Order Thinking Skills (HOTS) outlined in Bloom's revised taxonomy include the advanced cognitive competencies of analysis (C4), evaluation (C5), and creation (C6) (Kwangmuang et al., 2021). These skills play an important role in fostering critical and creative thinking in students (Wu et al., 2025), empowering them to navigate complex real-world problems effectively (Nurherdiati, 2022). Engaging students in activities aligned with HOTS not only enhances analytical abilities but also deepens their understanding and application of knowledge across various subjects (Kim et al., 2020; Pujawan et al., 2022; Sutika et al., 2023). Furthermore, the integration of technology, such as mobile platforms and smart classrooms, has been shown to stimulate motivation and engagement in learning, thereby supporting the development of higher-order skills (Lu et al., 2021). As a result, educational methodologies that combine structured peer interaction and problem-solving strategies significantly improve cognitive skills essential for educational success (Indriani et al., 2024; Kenny Tirtayasa et al., 2022; Sukawati et al., 2023).

2.4. Deep Learning and Critical Reading Pedagogy

The deep learning pedagogical approach emphasises a constructivist framework, making critical thinking and problem-solving essential components of education (Amorim & Ribeiro-Silva, 2025). In the context of critical reading, the principles of deep learning are reflected through several key strategies: (1) encouraging readers to question the author's purpose and bias, which is in line with

deep learning's emphasis on questioning assumptions; (2) requiring students to synthesize information from multiple texts, reflecting deep learning's goal of making meaningful connections; and (3) promoting metacognitive awareness in which students monitor their own comprehension processes. Deep learning environments facilitate the kind of sustained engagement necessary to develop critical reading competencies (Clark & Post, 2021). This pedagogy encourages students to become active participants in the learning process (Ekinci et al., 2023), and constructivist strategies enhance students' ability to critically evaluate textual arguments (Yang et al., 2025). Furthermore, inquiry-based learning, closely related to constructivist strategies, offers rich opportunities for students to question, explore (Kim et al., 2020), and reflect, significantly enhancing problem-solving abilities (Hartini et al., 2023). The transition from passive to active learning environments aligns with the need for deeper understanding and critical reasoning, which are essential for today's learners (Afrilyasanti et al., 2023; Chipamaunga et al., 2025).

3. Methodology

3.1. Study Design

This study employs a Research and Development (R&D) design, based on the Borg & Gall (1983) model, to develop adaptive reading materials tailored to readability and Higher-Order Thinking Skills (HOTS)-oriented assessments for students in grades 4 to 12 across primary and secondary education. This model consists of ten stages, which were simplified into five main phases in this study: (a) Preliminary Survey and Needs Analysis; (b) Readability Analysis and Reconstruction; (c) Development of HOTS Assessment; (d) Multi-Party Validation; and (e) Field Effectiveness Test.

The research process was carried out through a sequence of interconnected stages to ensure comprehensive product development and validation:

- a. Preliminary Survey and Needs Analysis Stage: Identify reading texts and group them into seven thematic categories relevant to the national curriculum.
- b. Readability Analysis and Reconstruction Stage: Analysis and adjustment of texts using the Flesch Reading Ease (FRE) and Flesch–Kincaid Grade Level (FKGL) formulas to align complexity with grade level.
- c. HOTS Assessment Development Stage: Design ten HOTS-oriented assessment questions (C3–C6) for each grade level, based on the reconstructed texts.
- d. Multi-Stakeholder Validation Stage: A validation process involving students, teachers, and experts to test the product's feasibility.
- e. Field Effectiveness Test (Experiment) Stage: Conducting trials in selected schools using a quasi-experimental design (Pre-Test and Post-Test) to prove the product's impact on students' critical reading competence.

3.2. Participants

The research participants consisted of three main groups involved in the validation process, ensuring a comprehensive assessment of the product from various perspectives:

- a. Students (N≈3000): Acted as end users and were involved in testing the readability and difficulty level of the material.
- b. Teachers (N≈250): Evaluate curricular relevance and classroom usability.
- c. Experts (N=12): Specialists in education and linguistics who assess content validity, linguistic accuracy, and pedagogical coherence.

d. Effectiveness Test Population: A sample of students (N=30) from a specific grade level was involved in the Pre-Test and Post-Test stages to measure improvements in critical reading competence.

Participants were selected using a stratified random sampling method with the following structure: First, three geographical-cultural clusters in Central Java were identified (North Coast, Negarigung, and Banyumasan). Second, within each cluster, schools were differentiated based on type (public/private) and level (elementary/junior high/senior high), with proportional allocation to ensure representation across all nine grade levels (4-12). Third, from each selected school, classes were randomly selected, and all students in those classes participated. This stratification ensures geographical diversity (33.3% per cluster), a balance of school types (60% public, 40% private), and balanced grade level representation (approximately 333 students per grade).

3.3. Data Collection Tools

The instruments used were designed to collect objective data (readability measurements), cognitive data (HOTS performance), and perception data (feasibility validation):

a. Readability Formula: The Flesch Reading Ease (FRE) and Flesch–Kincaid Grade Level (FKGL) formulas were used as objective instruments to analyse and reconstruct texts, producing a mathematically measurable readability index.

b. HOTS-Oriented Assessment Items: Ten assessment items (C3–C6) per grade level are the main instrument for measuring critical reading competence, focusing on Analysis, Evaluation, Synthesis, and Creation.

c. Likert Scale: A 1–5 scale is used to collect quantitative data on the perceptions of students, teachers, and experts during validation. The aspects assessed include text readability, relevance of HOTS items, difficulty level, suitability for critical reading skills, usefulness in the classroom, and overall quality.

d. Open Feedback: Used to obtain in-depth qualitative insights from teachers and experts regarding strengths and areas for improvement.

3.4. Validity of Data Collection Tools

The validity of the instruments was ensured through a multi-dimensional process involving assessment by three stakeholder groups at the multi-stakeholder validation stage:

a. Content and Construct Validity: The HOTS assessment items were assessed by 12 experts to ensure that the intended C3 to C6 skills were measured.

b. Criterion Validity (Curricular and Pragmatic): Evaluated by 250 teachers regarding curricular relevance and classroom usability.

c. Psychometric Validity (New): Proven through statistical analysis of HOTS assessment items (Difficulty Index and Discrimination Index) to ensure their structural challenge and reliability.

3.5. Data Analysis Procedure

The data analysis procedure adopted a mixed methods approach, combining quantitative and qualitative analysis:

a. Quantitative Analysis:

1) Descriptive Statistics: Used to analyse the readability index (FRE/FKGL) and Likert Scale validation ratings (mean and percentage) to capture product acceptance trends.

2) Inferential Statistics (New): Used for effectiveness test data (Pre-Test and Post-Test). A paired-samples t-test was conducted to assess the significance of the increase in critical reading competence.

3) Item Quality Analysis (New): Calculate each HOTS assessment item's Difficulty Index and Discriminative Power.

b. Qualitative Analysis: Insights from open feedback from teachers and experts are analyzed thematically to identify strengths and areas for improvement.

3.6. Ethical Considerations

Ethical considerations were a priority throughout the study. Ethical measures implemented included securing Institutional Approval, obtaining Informed Consent from all participants, and ensuring the confidentiality of participant data to maintain the anonymity of responses.

4. Results

The research results are based on design data, validation, item quality testing, and effectiveness testing of the intervention.

4.1. Readability and Textual Complexity Progression Analysis

Confirmatory analysis demonstrates that the reconstruction stage successfully created measurable gradations of textual complexity appropriate for the progression through primary and secondary education levels.

Table 1. Readability Analysis of Reconstructed Materials (FRE and FKGL)

| Grade | Flesch Reading Ease (FRE) | Flesch–Kincaid Grade Level (FKGL) |
|-------|---------------------------|-----------------------------------|
| 4 | 85.2 | 3.2 |
| 5 | 78.5 | 4.1 |
| 6 | 72.1 | 5.0 |
| 7 | 65.8 | 6.2 |
| 8 | 61.5 | 7.3 |
| 9 | 58.4 | 8.0 |
| 10 | 54.3 | 9.1 |
| 11 | 50.2 | 10.3 |
| 12 | 45.1 | 11.5 |

The Flesch Reading Ease (FRE) score gradually decreased from 85.2 in Grade 4 to 45.1 in Grade 12, reflecting increased text difficulty. Simultaneously, the Flesch–Kincaid Grade Level (FKGL) score shows a linear increase from Grade Level 3.2 (Grade 4) to 11.5 (Grade 12). The overall average FRE is 63.0, which is classified as “easy to slightly difficult,” indicating a balance between accessibility and cognitive challenge.

4.2. Distribution and Calibration of HOTS Assessment Items

The HOTS assessment was deliberately calibrated to align with cognitive development according to the Revised Bloom's Taxonomy (C3 to C6). Each grade level has 10 assessment items.

Table 2. Distribution of HOTS-Oriented Items (Bloom's Taxonomy)

| Grade | Total Items | C3 (Analysis) | C4 (Evaluation) | C5 (Synthesis) | C6 (Creation) |
|-------|-------------|---------------|-----------------|----------------|---------------|
| 4 | 10 | 3 | 2 | 3 | 2 |
| 5 | 10 | 2 | 3 | 3 | 2 |
| 6 | 10 | 2 | 3 | 4 | 1 |
| 7 | 10 | 3 | 2 | 3 | 2 |
| 8 | 10 | 3 | 2 | 4 | 1 |
| 9 | 10 | 4 | 2 | 3 | 1 |
| 10 | 10 | 3 | 3 | 2 | 2 |
| 11 | 10 | 4 | 3 | 2 | 1 |
| 12 | 10 | 4 | 3 | 2 | 1 |

The distribution of test items shows consistent progression in cognitive demands. Lower grades place strong emphasis on Analysis (C3) and Evaluation (C4), while higher grades integrate more complex Synthesis (C5) and Creation (C6) tasks.

4.3. Results of Multi-Stakeholder Product Feasibility and Acceptance Validation

Validation results, using a Likert scale (1–5), show exceptionally high acceptance across all participant groups.

Table 3. Results of Adaptive Material and HOTS Item Validation (Likert Scale 1–5)

| Aspect | Students (N≈3000) | Teachers (N≈250) | Experts (N=12) |
|---|----------------------|---------------------|-------------------|
| Text Readability | 4.3 | 4.5 | 4.6 |
| Hots Item Relevance | 4.5 | 4.7 | 4.8 |
| Level Of Difficulty | 4.2 | – | – |
| Suitability For Critical Reading Skills | 4.4 | – | – |
| Usefulness In The Classroom | – | 4.6 | – |
| Overall Quality | – | – | 4.7 |

All aspects achieved an average score above 4.0/5.0. This indicates a strong consensus regarding the product's quality and suitability. Teachers gave high scores for Classroom Usefulness (4.6), and experts gave the highest scores for HOTS Item Relevance (4.8), validating the challenges and usability of the product in the field.

4.4. Psychometric Analysis of HOTS Item Quality

Statistical analysis of the assessment items provides empirical evidence of the instrument's structural validity.

Table 4. Results of the HOTS Assessment Psychometric Item Analysis (Grade 8, N=30)

| HOTS Level (C) | Item ID | Difficulty Index (p) | Discrimination Power (D) | Functional Quality |
|-----------------|---------|--------------------------|------------------------------|--|
| C3 (Analysis) | HOTS_01 | 0.65 | 0.45 | Good (Optimal p , D strong) |
| C3 (Analysis) | HOTS_02 | 0.72 | 0.38 | Acceptable (Slightly easy, D strong) |
| C3 (Analysis) | HOTS_03 | 0.55 | 0.41 | Good (Optimal p , D strong) |
| C4 (Evaluation) | HOTS_04 | 0.48 | 0.35 | Good (Challenging, D good) |
| C4 (Evaluation) | HOTS_05 | 0.60 | 0.33 | Good (Optimal p , D good) |
| C5 (Synthesis) | HOTS_06 | 0.35 | 0.28 | Acceptable (Challenging, D sufficient) |
| C5 (Synthesis) | HOTS_07 | 0.41 | 0.31 | Good (Challenging, D good) |
| C5 (Synthesis) | HOTS_08 | 0.38 | 0.25 | Acceptable (Challenging, D adequate) |
| C5 (Synthesis) | HOTS_09 | 0.50 | 0.30 | Good (Optimal p , D good) |
| C6 (Creation) | HOTS_10 | 0.25 | 0.18 | Needs Revision (Too difficult, D weak) |

Most items (C3 to C5) showed an optimal Difficulty Index ($p = 0.30$ – 0.70) and a strong Discrimination Index ($D > 0.20$). However, item C6 (Creation) showed weak discrimination ($D = 0.18$) and excessive difficulty ($p = 0.25$). Analysis shows that the item's complexity exceeds students' developmental readiness, particularly by requiring them to think creatively and evaluatively simultaneously. The item has been revised by: (1) simplifying the task structure while maintaining cognitive demands; (2) providing clearer guidance in the instructions; and (3) adjusting the rubric to capture partial creative responses better. A pilot test of the revised item ($n=15$) resulted in improved metrics ($D = 0.31$, $p = 0.38$).

4.5. Intervention Effectiveness Test (Pre-Test and Post-Test of Critical Reading Competence)

A Paired Samples T-Test was conducted on the critical reading competency scores of 30 students ($n=30$) to assess the product's impact on students' critical reading competency.

Table 5. Effectiveness Test Results: Descriptive and Inferential Statistics of Critical Reading Competency (N=30)

| Parameter | Mean (Rata-Rata) | Std. Deviation (SD) | Std. Error Mean |
|-------------------|------------------|---------------------|-------------------------|
| Pre-Test Score | 62.97 | 5.774 | 1.054 |
| Post-Test Score | 75.53 | 5.144 | 0.939 |
| Paired Difference | -12.567 | 0.935 | 0.171 |
| t-statistic | -73.595 | df=29 | Sig. (2-tailed) = 0.000 |

Average Increase: There was a clear increase in the average critical reading competency score, from 62.97 (pre-test) to 75.53 (post-test), with an average difference of 12.567 points. The two-tailed significance value (Sig. 2-tailed) is 0.000. Since this value is much smaller than the standard significance threshold ($\alpha=0.05$), the null hypothesis (that there is no difference between the scores) is rejected.

These results prove that the intervention using adaptive reading materials based on readability and HOTS assessment significantly improves students' Critical Reading Competence. The very strong correlation coefficient ($r = 0.992$, $p = 0.000$) indicates that the Pre-Test and Post-Test scores are closely related, supporting the validity of the paired experimental design.

5. Discussion

5.1. Adaptive Design and Progression of Textual Complexity

This discussion focuses on the study's success in designing linguistic scaffolding for Critical Reading Competence. The finding that measured textual complexity (as assessed by readability formulas) increases linearly across grade levels validates the methodological assumption that critical reading competence must be developed through calibrated text challenges. In Indonesia, where teaching materials are often static, this adaptive approach ensures that students are exposed to cognitive input that is neither too easy nor too difficult, an important prerequisite for fostering higher-order thinking. Previous research has shown that educational resources often exceed the recommended readability level, making it difficult for students to engage with the material effectively (Hartnett et al., 2025; Sahhar et al., 2025). The readability variable control used in this study, based on the Flesch-Kincaid, effectively creates such a scaffolding environment, in line with the ongoing call to improve readability in educational resources to ensure they are understandable to all learners (Palupi, 2024). Thus, this study contributes to the R&D literature by providing a validated model for integrating objective readability as a basis for curriculum design.

5.2. Product Validity and Assessment Challenges

High validation across stakeholder groups—students, teachers, and experts reveals an important finding: the materials successfully balance theoretical rigor with practical applicability. Teachers' high ratings of classroom usefulness (4.6) indicate that texts adapted for readability meet real pedagogical needs, while experts' support for the relevance of HOTS questions (4.8) confirms alignment with cognitive development theory. This dual validation from practitioners and theorists reinforces the argument that effective critical reading instruction requires accessible content and cognitively challenging assessments, suggesting that these assessments accurately measure higher-order thinking skills (Ceballos et al., 2026). Previous research has shown that the development of HOTS instruments requires careful consideration to ensure that critical dimensions (Simons et al., 2020), problem solving, and creation are truly measurable (Mukhlis et al., 2023). The strong performance of

items C3, C4, and C5 provides evidence of structural validity that complements the content validity that has been validated by experts, in line with the importance of using instruments developed using standard steps, so that the results can be recognized as valid (Foo, 2025).

5.3. Effectiveness of Interventions in Optimizing Critical Reading Competence

The effectiveness test findings prove that interventions using adaptive materials and HOTS assessments result in a statistically significant increase in students' Critical Reading Competence. This empirical evidence is crucial, as it shifts the focus of studies from merely “developing” to “optimizing” critical reading competency. The significant improvement indicates that an integrated approach—providing texts of calibrated complexity followed by tasks that demand higher-order thinking—is highly effective. These findings are in line with previous research confirming that integrating HOTS into reading instruction fosters cognitive growth and student engagement (Istifadah et al., 2023), and that it goes beyond basic comprehension to support interpretation, evaluation, and decision-making about the material being read (Begimbetova et al., 2025). Furthermore, other studies support that fostering HOTS requires the integration of active learning techniques that stimulate critical thinking and questioning among students (Adiredja et al., 2023), which is a natural result of using HOTS assessments. This approach, which is also in line with the use of project-based learning in improving critical thinking and problem-solving skills (Afriani, 2022), proves that this product can encourage critical and creative thinking in students (Aryana et al., 2025). By quantitatively demonstrating its effectiveness, this study supports broader recommendations for educators to systematically adopt HOTS-based strategies to causally improve students' critical reading competencies.

6. Conclusion

This Research and Development Study validated adaptive reading materials based on readability and HOTS-oriented assessment at the Indonesian elementary and secondary education level. This study comprehensively proves: (1) the success of linguistic scaffolding design through FKGL and FRE calibration; (2) high acceptance and product feasibility by multi-stakeholders; and most importantly, (3) the statistically significant effectiveness of the intervention in improving students' Critical Reading Competence. With valid evidence and empirical impact, this product represents a significant step forward in national efforts to improve literacy and higher-order thinking skills.

7. Limitations and future research

The main limitation of this study lies in the effectiveness testing phase. Although the pre-test and post-test designs demonstrated a significant increase in KMK, this study did not include a control group to compare the product's impact with conventional teaching methods. This limits the ability to isolate the absolute causal effect and claim superiority over other interventions. Another limitation is that, although the psychometric analysis was conducted, it identified weaknesses in the Creation item (C6), suggesting that the measurement instrument may not be fully valid for the highest-level thinking dimension. Therefore, further research should focus on: (1) Conducting a randomised controlled trial (RCT) comparing the intervention group and the control group to strengthen the evidence of effectiveness; and (2) Revising and recalibrating the C6 assessment item to achieve an optimal Discrimination Index, followed by further trials to ensure a perfect assessment challenge.

8. Suggestion

The practical implications of this study are substantial for the national education system. This R&D product, which has been proven valid, feasible, and effective, offers a tool ready for use by primary and secondary education teachers to transition from recall-based reading instruction to critical reading. Practical suggestions include: First, extensive teacher professional development should focus on integrating this adaptive material into the daily curriculum and on using readability indices to guide scaffolding. Second is Curriculum Policy Adoption, in which educational institutions

can consider introducing this adaptive design model across all curricular reading materials to ensure consistent cognitive progression for students. Third, Scaling-Up, which is the development of this product into an interactive digital platform that can automatically adjust the text's readability level (FRE/FKGL) based on students' real-time performance on their HOTS items, leading to true personalised learning.

Declarations

Author Contributions. S.: Conceptualization, literature review, and initial draft writing. M.B.S.: Methodology, data collection, and data analysis. L.V.P.: Data analysis and validation of findings. E.K.: Review and refinement of content. S.S.: Review-editing and final manuscript writing. All authors have read and approved the final version of the article to be published.

Conflicts of Interest. The authors declare that they have no conflicts of interest in publishing the results of this study.

Data Availability Statement. The data supporting the findings of this study are available from the corresponding author upon reasonable request.

References

- Adam, P., Kleemans, T., Droop, M., Amorim, A. N., & Segers, E. (2025). A digital boost of the home literacy environment positively impacts early reading and vocabulary development of Indonesian 1 1st graders. *The Journal of Educational Research*, 118(5), 425–436. <https://doi.org/10.1080/00220671.2025.2490963>
- Adiredja, R. K., Hartati, T., & Riyana, C. (2023). Development of Integrated Writing Materials Based on Multiliteracies and High-Order Thinking Skills. *Jurnal Kependidikan: Jurnal Hasil Penelitian Dan Kajian Kepustakaan Di Bidang Pendidikan, Pengajaran Dan Pembelajaran*, 9(3), 882. <https://doi.org/10.33394/jk.v9i3.8757>
- Afriani, N. R. (2022). Effectiveness Of The Problem Based Learning Model On Critical Thinking Ability About Science Subject For Fifth Grade Elementary School Students. *EduHumaniora | Jurnal Pendidikan Dasar Kampus Cibiru*, 14(1), 46–58. <https://doi.org/10.17509/eh.v14i1.36791>
- Afrilyasanti, R., Basthomi, Y., & Zen, E. L. (2023). EFL students' participations and teachers' roles in online discussion forum for critical media literacy learning. *Contemporary Educational Technology*, 15(2), ep414. <https://doi.org/10.30935/cedtech/12965>
- Amorim, C., & Ribeiro-Silva, E. (2025). Preservice physical education teachers' perceptions of initial teacher education. *European Physical Education Review*, 31(3), 355–373. <https://doi.org/10.1177/1356336X241276825>
- Ardhian, T., Ummah, I., Anafiah, S., & Rachmadtullah, R. (2020). Reading and Critical Thinking Techniques on Understanding Reading Skills for Early Grade Students in Elementary School. *International Journal of Instruction*, 13(2), 107–118. <https://doi.org/10.29333/iji.2020.1328a>
- Aryana, S., Sobari, T., Wikanengsih, Fauziya, D. S., & Mustika, R. I. (2025). The Effectiveness of 21st Century Competency Assessment Model and HOTS in Improving Writing Skills: An Experimental Research. *Journal of Language Teaching and Research*, 16(4), 1115–1128. <https://doi.org/10.17507/jltr.1604.07>
- Badolo, M., Malik, M. A., Nur, R., & Latifa, A. (2025). The Impact of Metacognitive Strategy Training on Higher-Order Thinking Skills (HOTS) in High School Mathematics: A Quasi-Experimental Study. *International Journal of Environment, Engineering and Education*, 7(2), 146–157. <https://doi.org/10.55151/ijeedu.v7i2.302>

- Barton, G., Riddle, S., & Lowien, N. (2024). ““The last bastion of democracy””: teachers’ perceptions of the democratic potential of English curriculum”. *English Teaching: Practice & Critique*, 23(4), 475–491. <https://doi.org/10.1108/ETPC-08-2024-0127>
- Beck, K. B., MacKenziem, K. T., Kirby, A. V., McDonald, K., Moura, I., Breitenfeldt, K., Rutenberg, E., Kumar, T., Mancino, J., Sabatello, M., Roth, S., & Nicolaidis, C. (2025). Guidelines for Creating Accessible Consent Materials and Procedures: Lessons from Research with Autistic People and People with Intellectual Disability. *Autism in Adulthood*. <https://doi.org/10.1089/aut.2024.0263>
- Begimbetova, G. A., Retnawati, H., Kassymova, G. K., Triyono, M. B., Azman, M. N. A., & Sodikin, R. A. (2025). Developing Students’ HOTS Digital Literacy Profile by Using IRT. *Journal of Curriculum Studies Research*, 7(2), 155–176. <https://doi.org/10.46303/jcsr.2025.16>
- Belda-Medina, J. (2024). Inclusive Education through Digital Comic Creation in Higher Learning Environments. *Social Sciences*, 13(5), 272. <https://doi.org/10.3390/socsci13050272>
- Burke, E., Harkins, P., & Arumugasamy, M. (2024). A Critical Assessment of the Quality and Readability of Information on Gastric Cancer Available on the Internet for Patients. *Cureus*. <https://doi.org/10.7759/cureus.71691>
- Ceballos, H., Bogaart, T. van de, van Ginkel, S., Spandaw, J., & Drijvers, P. (2026). How collaborative problem solving promotes higher-order thinking skills: A systematic review of design features and processes. *Thinking Skills and Creativity*, 59, 102001. <https://doi.org/10.1016/j.tsc.2025.102001>
- Chipamaunga, S. R., Prozesky, D. R., Kafumukache, E., Katowa-Mukwato, P., Dithole, K. S., van der Merwe, L. J., Kagawa, M. N., & Gwini, R. (2025). Needs for Technology-Enhanced Health Professions Education in Eastern and Southern Africa: Protocol for a Descriptive, Cross-Sectional Survey. *JMIR Research Protocols*, 14, e67331. <https://doi.org/10.2196/67331>
- Clark, C. E. J., & Post, G. (2021). Preparation and synchronous participation improve student performance in a blended learning experience. *Australasian Journal of Educational Technology*, 187–199. <https://doi.org/10.14742/ajet.6811>
- Ekinci, F., Bektaş, O., Karaca, M., & Yiğit, K. N. (2023). The flipped learning perception scale: A validity and reliability study. *Education and Information Technologies*, 28(11), 14139–14166. <https://doi.org/10.1007/s10639-023-11731-7>
- Foo, S. Y. (2025). Investigating gifted students’ higher-order thinking skills in a differentiated learning environment: A case study. *Gifted Education International*, 41(2), 236–259. <https://doi.org/10.1177/02614294241305766>
- Gil-Rojas, B., Castellanos-Alvarenga, L. M., Gaitán-Angulo, M., Gómez-Cacedo, M. I., & Pérez-Pertuz, J. (2025). Critical Reading and Human Rights in Police Students: Development of Citizen Competencies to Prevent Crimes. *SAGE Open*, 15(2). <https://doi.org/10.1177/21582440251335847>
- Hartini, S., Liliyasi, L., & Sinaga, P. (2023). Students’ problem solving skill in nuclear physics course through NPIRL. *Momentum: Physics Education Journal*, 7(2), 279–289. <https://doi.org/10.21067/mpej.v7i2.8567>
- Hartnett, D. A., Philips, A. P., Daniels, A. H., & Blankenhorn, B. D. (2025). Readability of Online Foot and Ankle Surgery Patient Education Materials. *Foot & Ankle Specialist*, 18(1), 9–18. <https://doi.org/10.1177/19386400221116463>

- Indriani, R. P., Kurniati, T. H., & Ristanto, R. H. (2024). Implementing ARVi media to enhance students high order cognitive skills. *JPBI (Jurnal Pendidikan Biologi Indonesia)*, 10(2), 572–579. <https://doi.org/10.22219/jpbi.v10i2.32446>
- Istifadah, I., Usriyah, L., & Rahmawati, R. (2023). Triggering Interest in Learning through Environmentally Based Space Management. *Al-Tanzim: Jurnal Manajemen Pendidikan Islam*, 7(4), 1161–1176. <https://doi.org/10.33650/al-tanzim.v7i4.6841>
- Kenny Tirtayasa, I. K., Nitiasih, P. K., & Budiarta, L. G. R. (2022). HOTS-Based Worksheet in English Learning Activity for Eight Grade Students of Junior High School. *Jurnal Pendidikan Bahasa Inggris Undiksha*, 10(1), 73–81. <https://doi.org/10.23887/jpbi.v10i1.49824>
- Kim, H. J., Yi, P., & Hong, J. I. (2020). Students' Academic Use of Mobile Technology and Higher-Order Thinking Skills: The Role of Active Engagement. *Education Sciences*, 10(3), 47. <https://doi.org/10.3390/educsci10030047>
- Kwangmuang, P., Jarutkamolpong, S., Sangboonraung, W., & Daungtod, S. (2021). The development of learning innovation to enhance higher-order thinking skills for students in Thailand junior high schools. *Heliyon*, 7(6), e07309. <https://doi.org/10.1016/j.heliyon.2021.e07309>
- Li, J. W., Leung, S. K. Y., Wang, M. D., & Yuen, M. (2024). Developing Talented Children's Computational Thinking Through Multimodal Literacies in Pop-Up Storybooks: A Case Study in Hong Kong. *Education Sciences*, 14(12), 1377. <https://doi.org/10.3390/educsci14121377>
- Lu, K., Yang, H. H., Shi, Y., & Wang, X. (2021). Examining the key influencing factors on college students' higher-order thinking skills in the smart classroom environment. *International Journal of Educational Technology in Higher Education*, 18(1), 1. <https://doi.org/10.1186/s41239-020-00238-7>
- M. Mujiya Ulkhaq. (2024). Factors Influencing Indonesian Students' Performance on PISA 2018. *Proceeding of The International Conference of Inovation, Science, Technology, Education, Children, and Health*, 4(1), 74–87. <https://doi.org/10.62951/icistech.v4i1.78>
- Mardiana, N., Maulina, B., Mardiani, N., Sabar, S., & Collantes, L. M. (2025). Enhancing Critical Thinking Skills Through Android-Assisted Virtual Physics Learning: A Focus On HOTS Development. *Jurnal Ilmiah Ilmu Terapan Universitas Jambi*, 9(2), 513–526. <https://doi.org/10.22437/jiituj.v9i2.34373>
- Maxnun, L., Kristiani, K., & Sulistyningrum, C. D. (2024). Development of hots-based cognitive assessment instruments: ADDIE model. *Journal of Education and Learning (EduLearn)*, 18(2), 489–498. <https://doi.org/10.11591/edulearn.v18i2.21079>
- Mukhlis, M., Suwandi, S., Rohmadi, M., & Setiawan, B. (2023). Higher Order Thinking Skills in Reading Literacy Questions at Vocational High Schools in Indonesia. *International Journal of Language Education*, 7(4). <https://doi.org/10.26858/ijole.v7i4.37603>
- Nurherdiati, R. (2022). An Instructional Design to Improve Students' Higher Order Thinking Skills in Linear Programming. *J-PiMat: Jurnal Pendidikan Matematika*, 4(1), 459–465. <https://doi.org/10.31932/j-pimat.v4i1.1619>
- Palupi, T. M. (2024). Developing a learning management system for critical literacy course: A need analysis. *Bahasa Dan Seni: Jurnal Bahasa, Sastra, Seni, Dan Pengajarannya*, 51(1). <https://doi.org/10.17977/um015v51i12023p91>

- Panigrahi, A. R., & Bakshi, S. G. (2021). Preparation, validation, and evaluation of an information leaflet for patients undergoing day-care surgeries under general anesthesia at a busy tertiary care hospital. *Journal of Anaesthesiology Clinical Pharmacology*, 37(2), 243–248. https://doi.org/10.4103/joacp.JOACP_264_19
- Pujawan, I. G. N., Rediani, N. N., Antara, I. G. W. S., Putri, N. N. C. A., & Bayu, G. W. (2022). Revised Bloom Taxonomy-Oriented Learning Activities to Develop Scientific Literacy and Creative Thinking Skills. *Jurnal Pendidikan IPA Indonesia*, 11(1), 47–60. <https://doi.org/10.15294/jpii.v11i1.34628>
- Puspitadani, E., Yudea, F., & Loo, F. (2022). Educational Leadership and Learning Quality: The Influence of the Principal's Leadership Style on Teacher Performance. *Jurnal Ilmu Pendidikan Dan Humaniora*, 11(3), 206–220. <https://doi.org/10.35335/jiph.v11i3.24>
- Rahmawati, T., Wahyuningrum H, M. M., Bustari, M., Lestari, S., Ernawati, R. D., & Selviana, S. (2023). Science Learning Management Based on Higher Order Thinking Skills (HOTS). *Jurnal Penelitian Pendidikan IPA*, 9(1), 533–541. <https://doi.org/10.29303/jppipa.v9i1.3152>
- Randell, R. L., Wilson, H. P., Ragavan, M. I., Collins, A. B., Vail, J., Ramirez, S., Amodei, J., Mickiewicz, E., Krieger, M. S., Macon, E. C., & Hornik, C. P. (2025). Communicating Health Research With Plain Language. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 62. <https://doi.org/10.1177/00469580251357755>
- Sahhar, M., Singh, M., Mehta, T., Laud, A., Nassar, J. E., Farias, M. J., Diebo, B. G., & Daniels, A. H. (2025). Lost in Translation: Preoperative Orthopaedic Education Materials Significantly Exceed Recommended Reading Levels. *JBJS Open Access*, 10(3). <https://doi.org/10.2106/JBJS.OA.25.00143>
- Simons, M., Smits, T. F. H., & Janssenswillen, P. (2020). Newspapers as teaching tools for media literacy education: what makes teachers use newspapers in their classrooms? *Educational Media International*, 57(4), 332–352. <https://doi.org/10.1080/09523987.2020.1848510>
- Stepanenko, O., Suriak, M., Trubacheva, S., Sirenko, P., & Vlasiuk, I. (2023). Formation of Students' Reading Competence in the Conditions of Integrated Learning. *Cadernos de Educação, Tecnologia e Sociedade*, 16(3), 648–656. <https://doi.org/10.14571/brajets.v16.n3.648-656>
- Sukawati, F., Saude, S., Yusra, Y., & Ubadah, U. (2023). Inquiry Discovery Learning Method To Improve Student Learning Achievement In Islamic Religious Education Subjects. *International Journal Of Contemporary Islamic Education*, 5(2), 46–56. <https://doi.org/10.24239/ijcied.Vol5.Iss2.80>
- Sutika, I. M., Winaya, I. M. A., Rai, I. B., Sila, I. M., Sudiarta, I. N., Kartika, I. M., & Sujana, I. G. (2023). The Effectiveness of the Problem-Based Learning Model in Improving Higher-Order Thinking Skills and Character of Elementary School Students. *Jurnal Pendidikan Dan Pengajaran*, 55(3), 688–702. <https://doi.org/10.23887/jpp.v55i3.57636>
- Weng, T. (2023). Creating Critical Literacy Praxis: Bridging the Gap between Theory and Practice. *RELC Journal*, 54(1), 197–207. <https://doi.org/10.1177/0033688220982665>
- Wu, T., Sarwono, E., & Huang, Y. (2025). Incorporating Computational Thinking Into Virtual Laboratories to Enhance Learning Motivation, Engagement, and Higher-Order Thinking Skills. *Journal of Computer Assisted Learning*, 41(2). <https://doi.org/10.1111/jcal.70017>
- Yang, Y., Li, Y., Fu, J., Guo, D., & Xue, J. (2025). Effectiveness of a Four-Stage Death Education Model Based on Constructivist Learning Theory for Trainee Nursing Students. *Journal of Multidisciplinary Healthcare*, Volume 18, 1371–1380. <https://doi.org/10.2147/JMDH.S500169>

Yüceer-Çetiner, E., Kazan, D., Nesiri, M., & Basa, S. (2025). Evaluating the Competence of AI Chatbots in Answering Patient-Oriented Frequently Asked Questions on Orthognathic Surgery. *Healthcare*, 13(17), 2114. <https://doi.org/10.3390/healthcare13172114>

About the Contributor(s)

Prof. Subyantoro, is a professor of psycholinguistics. He actively carries out the Tri Dharma Perguruan Tinggi (Three Pillars of Higher Education), which includes developing lecture materials related to psycholinguistics, language acquisition and learning, language skills development, and research methodologies for Indonesian language and literature. He also conducts various linguistic research, training, and seminars for the professional development of educators. Armed with his knowledge of psycholinguistics, he has developed new courses such as Speech Therapy, Neurolinguistics, and Forensic Linguistics. Prof. Subyantoro serves as a Language Expert Witness at the Central Java Regional Police and focuses on the scientific field of Forensic Linguistics. For his contributions to the world of education, he was awarded the “Satria Educational Brand Award 2024”.

Email: bintoro@mail.unnes.ac.id

ORCID: <https://orcid.org/0000-0002-4365-5993>

Muhammad Badrus Siroj, is a lecturer in the Indonesian Language and Literature Education Study Program (S1), Faculty of Language and Arts, Universitas Negeri Semarang.

Email: badrussiroj@mail.unnes.ac.id

ORCID: <https://orcid.org/0000-0002-5446-1481>

Lisa Virdinarti Putra, Dr, is a lecturer in the Graduate Program of Educational Management at Universitas Ngudi Waluyo. She has research expertise in educational management and elementary school education.

Email: lisavirdinartiputra@gmail.com

ORCID: <https://orcid.org/0000-0001-8874-575X>

Endang Kusrini, is affiliated with the Faculty of Language and Arts, Universitas Negeri Semarang.

Email: endang@gmail.com

ORCID: <https://orcid.org/0009-0003-6340-5734>

Sisca Septiani, Dr, is a lecturer at Universitas Ngudi Waluyo, Semarang. Her research interests include educational management and education. She is also active in book writing, reviewing, and editing for Scopus-indexed international journals.

Email: sisseptiani@gmail.com

ORCID: <https://orcid.org/0000-0001-7806-0458>

Publisher's Note: *The opinions, statements, and data presented in all publications are solely those of the individual author(s) and contributors and do not reflect the views of Universitepark, EDUPIJ, and/or the editor(s). Universitepark, the Journal, and/or the editor(s) accept no responsibility for any harm or damage to persons or property arising from the use of ideas, methods, instructions, or products mentioned in the content.*
