Jurnal Teknologi Pendidikan, Desember 2025, 27 (3), 824-841
DOI: http://dx.doi.org/10.21009/JTP2001.6
p-ISSN: 1411-2744 e-ISSN: 2620-3081
Accredited by Directorate General of Strengthening for Research and Development
Available online at http://journal.unj.ac.id/unj/index.php/jtp

Innovative Design of Learning Evaluation Textbooks Based on The Merdeka Curriculum: An Effort to Optimize The Learning Process

A. Zuhdi[1], Rasmuin[2], Yuniar Setyo Marandy[3], Nuril Nuzulia[4], Novia Solichah[5(*)]
[1,2,3]Department of Islamic Education, Faculty Tarbiya and Teachers Training, UIN Maulana Malik Ibrahim, Malang, Indonesia
[4]Department of Education for Islamic Elementary School Teachers, Faculty Tarbiya and Teachers Training, UIN Maulana Malik Ibrahim, Malang, Indonesia
[5]Department of Psychology, Faculty of Psychology, UIN Maulana Malik Ibrahim, Malang, Indonesia

Received : October 21, 2025
Revised : December 9, 2025
Accepted : December 24, 2025

Abstract

The Merdeka Curriculum emphasizes competency-based and authentic assessment; however, no evaluation textbook has been developed to support these principles for higher-education students, particularly in teacher education programs. This study aimed to develop and validate a Merdeka Curriculum–based learning evaluation textbook to strengthen students' assessment competence. This Research and Development (R&D) study adapted seven steps of the Borg & Gall model. Participants included 30 sixth-semester Islamic Education students, three material experts, one design expert, and one practitioner. Data were collected through expert validation questionnaires, student response surveys, observations, and pretest–posttest assessments. Quantitative data were analyzed using paired samples t-test and effect size (Cohen's d). The textbook demonstrated high validity, with material expert (82.4%), design expert (84%), and practitioner (84%) ratings categorized as very feasible. Student attractiveness responses reached 85.3%. The pretest–posttest comparison showed a significant improvement ($M_{pre} = 72.00$; $M_{post} = 88.67$; $t = -13.20$; $p < 0.001$). The effect size was very large (Cohen's $d = 2.57$), indicating substantial gains in students' evaluation competence. The developed textbook is valid, practical, and effective for enhancing students' understanding and skills in learning evaluation aligned with the Merdeka Curriculum. Future development should include broader samples, multi-institution testing, and digital or interactive formats to increase scalability.

Keywords: Assessment Literacy; Textbook Development; Merdeka Curriculum; Curriculum-based Evaluation; Instructional Design

(*) Corresponding Author: noviasolichah@uin-malang.ac.id

How to Cite: Zuhdi, A., Rasmuin, R., Setyo Marandy, Y., Nuzulia, N., & Solichah, N. (2025). Innovative Design of Learning Evaluation Textbooks Based on The Independent Curriculum: An Effort to Optimize The Learning Process. *JTP - Jurnal Teknologi Pendidikan*, *27*(3), 824–841. https://doi.org/10.21009/jtp.v27i3.61189

INTRODUCTION

One of the competencies that teachers must master is pedagogical competence or teaching ability, and its sub-competencies include the ability to assess learning. In the "Learning Evaluation" course, students will understand the basic concepts of assessment, including the types of evaluation (formative, summative, diagnostic, etc.) and the objectives of each type of assessment. This course basically teaches

the basic concepts, theories, principles, and procedures of learning assessment. The weight of the course is 2 credits. Because the nature of this course is compulsory, students are required to complete and pass, and is a prerequisite for basic teaching competencies and field work practice. To support these competencies, learning strategies grounded in educational research are essential.

To optimize the learning process, several strategies and methods can be employed based on research findings from various academic studies. Although these strategies come from different areas of learning science, they all highlight the importance of structured learning support, which is also the foundation of developing effective textbooks. Here are some methods: Spacing and Retrieval Practice (Black & Allen, 2019; Carpenter et al., 2022); Metacognition (Carpenter et al., 2022); Generative Learning Activities (Fiorella & Mayer, 2015); Multimedia Learning Principles (Castro-Alonso et al., 2021); AI and Machine Learning (Han, 2024; Wang, 2025); Active Learning (Fei et al., 2018); Formative Feedback (Black & Allen, 2019); Quantitative Learning Strategies (Zhao et al., 2018); Online Education Optimization (Jiménez-Macías et al., 2024). These studies collectively show that well-designed learning materials—whether digital, technological, or conventional—play a crucial role in supporting student understanding. This reinforces the need for a structured and systematic textbook as proposed in this research. In relation to these strategies, the learning evaluation process also plays an important role in shaping prospective teachers' competencies.

In the learning evaluation process, prospective teachers are trained to design, implement and analyze assessments. They learn how to create valid and reliable assessment instruments, so that they can objectively evaluate student achievement (R. Febrianto & Puspitaningsih, 2020). In the context of the Merdeka Curriculum, learning evaluation must be oriented towards developing student competence and character. This course provides an understanding of how to apply curriculum principles in assessment practices.

Prospective teachers learn how to provide useful feedback to students, so that students can understand their strengths and areas for improvement, as well as plan improvement steps (Sukriadi et al., 2015). Through student-involved evaluation (e.g., self-assessment), this course teaches prospective teachers to encourage students to be more active in the learning process and take responsibility for their progress. Good assessment can improve the overall quality of learning. Teacher candidates who are skilled in evaluation can identify the most effective teaching strategies based on assessment results. This course prepares prospective teachers to adapt to various learning contexts and diverse student needs, including how to customize assessments for students with special needs. By understanding evaluation, prospective teachers can become more professional and confident in carrying out their duties as educators, as well as contribute to the overall improvement of the education system. Given the importance of evaluation, the role of textbooks as structured learning resources becomes increasingly relevant.

The existence of textbooks as media or learning resources for students during lectures basically helps facilitate learning. Together, the learning process is known to occur both directly and indirectly. Direct learning means that students interact with the instructor. Conversely, indirect learning means that students actively interact with media and other learning resources to support their education. Zaini

(2006) states that the media facilitates the interaction process between teachers and students and helps students learn effectively and optimally. However, despite their importance, current practices show several challenges related to the availability of appropriate evaluation textbooks.

Based on a preliminary survey of students taking learning evaluation courses from all study programs in the Faculty of Tarbiyah Science, the average student has difficulty in designing and implementing appropriate evaluations, due to the lack of clear guidelines. This can hinder the effective implementation of the curriculum. In addition, the field shows that the absence of an evaluation book based on the Merdeka Curriculum reflects a number of challenges and needs that must be addressed. The results of interviews with students assume that an evaluation book based on the Merdeka Curriculum can encourage the use of innovative and diverse assessment methods. The absence of this book can prevent 6th semester students from exploring various evaluation strategies that are more interesting and effective. The absence of an evaluation book can hinder opportunities for prospective teachers or students to develop their skills and knowledge in assessment. This is essential for the improvement of professionalism and teaching effectiveness. To address these needs, several innovative approaches can be used in designing textbooks.

To design innovative learning evaluation textbooks based on the Merdeka Curriculum, there are several strategies can be derived Integrating local wisdom into the curriculum can significantly enhance students' understanding and academic achievement (Salmia et al., 2024). Differentiated learning strategies, which allow students to choose learning materials based on their interests and talents, can lead to more efficient learning outcomes (Hima et al., 2025). Implementing AR in textbooks can create immersive and engaging learning experiences (Ramli et al., 2024). Utilizing DBR for the development and evaluation of educational materials ensures that the design is grounded in real-world educational settings (Ramli et al., 2024). Encouraging collaboration and problem-based learning (PBL) can foster creativity and critical thinking (May et al., 2020).

Incorporating blended learning approaches and digital resources, such as e-books and online libraries, can support independent learning and provide students with access to a wider range of materials (Ruliah et al., 2020). Conducting regular evaluations to identify areas for improvement and making necessary adjustments can help in maintaining the effectiveness of the curriculum (P. T. Febrianto et al., 2025). Providing teachers with adequate training and resources to implement innovative teaching strategies is crucial. This includes workshops, professional development programs, and access to teaching platforms (Aris et al., 2025). Designing textbooks that support student-centered learning, where students take control of their learning process, can lead to better engagement and learning outcomes (Goldman & Pellegrino, 2015). Leveraging technology, such as web-based interactive learning systems and digital innovation teaching references, can enhance the learning experience and make it more accessible (Nintarat et al., 2023). These perspectives reinforce the urgency of developing a textbook aligned with the Merdeka Curriculum.

Based on these problems, a solution is needed. The findings indicate the need for collaboration between curriculum developers, lecturers, teachers, and educational institutions to develop relevant and applicable evaluation books. That

is, through the development of textbooks that are systematically designed according to the learning system design model so that the results are in accordance with internal needs. The purpose of the study was to develop a learning evaluation textbook that meets the needs of students based on the problems found, and to test the suitability of the developed textbook. As a result, it can meet the learning needs of students with textbook products that are suitable as media or learning resources (Prastowo, 2011). The importance of this development is also highlighted in relevant literature.

The importance of textbooks as media in the learning process can be seen from the student's point of view based on their features and benefits(Japa, 2012). (Japa, 2012) states that the selection and determination of appropriate learning media is based on needs, and the use of media is becoming increasingly urgent. Regarding the functions and benefits of textbooks, Prastowo (2013) states that the function of textbooks is for students to: (1) self-study; (2) learn anytime and anywhere; 3) self-study; (4) selected learning material; and (5) direct learning activities according to ability. The advantages of teaching materials include: (1) make learning more interesting, (2) provide opportunities for independent learning under student guidance, and (3) facilitate the acquisition of skills that need to be learned.

Research or development is an effort to develop and produce products in the form of materials, media, tools, and / or learning strategies for learning purposes in the classroom or laboratory (Fajri et al., 2023). The purpose of development is to improve the quality of learning processes and outcomes, both in terms of material content and methods or strategies. Surbakti (2017) stated that the development of textbooks used as media and/or learning resources can improve student outcomes. The same thing was also found by Marlina and Sholehun (2021) stated that learning outcomes can be significantly improved by using textbooks designed according to the needs and characteristics of students. Therefore, it can be concluded that textbook development is not perfunctory, but requires a systematic design related to the learning system, designing messages, methods or strategies, and paying attention to the characteristics of students and the results can be seen.

Based on the literature above, no studies have developed a learning evaluation textbook specifically aligned with the Merdeka Curriculum. Existing works discuss general textbook development or assessment concepts, but they do not address the need for evaluation materials that reflect the curriculum’s competency-based and authentic assessment principles. This gap shows that students still lack structured guidance for designing and implementing Merdeka Curriculum–oriented evaluations. Therefore, developing this textbook is necessary to provide relevant instructional support for FITK students, particularly in the Islamic Religious Education Study Program, and to strengthen their readiness in learning evaluation courses Merdeka Curriculum.

METHODS

1. Research Design

This study employed Research and Development (R&D) using an adapted

Borg and Gall model (Laws et al., 2013). This design was selected because it provides a systematic procedure for developing educational products such as textbooks, learning instruments, and evaluation tools. The Borg and Gall model originally consists of 10 steps (Aka, 2019); however, this study adopted only seven due to time and resource constraints. The adapted steps still follow the core logic of product development—needs analysis, product planning, product development, expert validation, revision, field testing, and final evaluation—ensuring methodological rigor while maintaining feasibility.

2. Population and Sample

The population of this study consisted of sixth-semester students of the Islamic Education Study Program at the Faculty of Tarbiyah and Teacher Training (FITK), UIN Maulana Malik Ibrahim Malang. From this population, a sample of 30 sixth-semester students was selected to participate in the field testing phase of the textbook implementation. In addition to student participants, several experts were involved in the validation process to ensure the quality and feasibility of the developed textbook.

Table 1. Population and Sample

Variable	Category	n	Percentage
Gender	Male	12	40%
	Female	18	60%
Age	20 years	8	26.7%
	21 years	15	50%
	22 years	7	23.3%
Semester	6th Semester	30	100%
GPA (Cumulative)	3.00–3.24	6	20%
	3.25–3.49	14	46.7%
	3.50–3.75	10	33.3%

3. Ethical Considerations

All participants provided informed consent prior to involvement in the study. Participation was voluntary, and students were assured that data would be used solely for research purposes. Expert validators were formally invited and agreed to participate as evaluators.

4. Development procedure

a. The first step is for researchers to identify problems and needs in the field. In the context of developing learning evaluation books, researchers will examine:
   1) Problems or difficulties faced by teachers in conducting learning evaluations.

2) Existing evaluation books or tools, as well as gaps or needs that have not been covered.
3) Literature review to find relevant evaluation theories and practices.

b. Step two, based on the results of the initial research, the researcher developed an evaluation book development plan that would meet the identified needs. Some things to consider in this planning include:
   1) Purpose of the evaluation book.
   2) Target users are 6th semester PAI students
   3) Structure and format of the book (e.g., type of evaluation, rubrics, instruments used).
   4) Content to be included in the book, such as evaluation questions, instructions for use, and ways of processing evaluation results.

c. The third product researchers began to develop the evaluation book in accordance with the planned design. The evaluation book contains various components, including:
   1) Evaluation questions: Test questions or assignments that can be used to assess student learning outcomes.
   2) Assessment rubric: Clear criteria for assessing the quality or success of students in achieving targeted competencies.
   3) Instructions and guidance: For lecturers or evaluators on how to use the book in learning evaluation.

d. In the fourth step, after the prototype of the evaluation book is completed, it is tested in small groups or in the form of validation by experts (such as education experts or evaluation experts) to assess the quality of the book. This validation aims to ensure that the content and structure of the book are in line with the objectives to be achieved and relevant to the learning context.

e. Step Five Based on the results of expert validation, researchers revised the evaluation book, both in terms of content, format, and presentation. These improvements aim to improve the quality of the book to make it more effective in the learning evaluation process.

f. In the sixth stage, the revised evaluation book is then tested in the field, either in small groups or with different groups of students and lecturers. At this stage, researchers collect data related to the use of the evaluation book, for example:
   1) Is the book easy for lecturers to use?
   2) Are the questions and assessment rubrics clear and understandable to students?
   3) Is the book effective in assessing student competence?

g. The seventh step is the last step of all the steps taken by the researcher. In this step, the researcher analyzes the data obtained from the field trial to assess the extent to which the evaluation book meets the desired objectives. Researchers also look for patterns or problems that arise during testing.

5. Test Run

1) Trial Design

This test is a stage to assess the effectiveness of the learning plan

coursebook based on the uniqueness of the curriculum developed. In this test, the developmental learning evaluation coursebook will be given to material experts and design experts, practitioners and 6th semester students of the Islamic Education Study Program, Faculty of Tarbiyah Sciences, UIN Maulana Malik Ibrahim Malang.

2) Test Subject

Experiments will be conducted with material experts, design experts, practitioners and field experiments with 6th semester students of the Islamic Education study program, Faculty of Tarbiyah Sciences, UIN Maulana Malik Ibrahim Malang.

a. Material Expert

The material experts comprised senior university lecturers with advanced expertise in assessment, Islamic Education, and educational evaluation. They possess extensive experience in developing and validating evaluation instruments, implementing curriculum-based assessment practices, and supervising students in evaluation-related coursework. Their qualifications ensured that the textbook content met standards of accuracy, relevance, and pedagogical appropriateness for higher education contexts. Category Qualifications of Content Experts:

1) Media proficiency
2) Have insight into the development of learning evaluation books Willingness to be a verifier of learning evaluation textbooks to improve the understanding and skills of 6th semester PAI students

b. Design expert

Instructors who became Design Expert Validators were selected from those who have the ability to study design. The design expert provides suggestions and feedback on the developed learning evaluation textbook.

c. Expert Practitioner

Skilled lecturers use learning evaluation textbooks to improve understanding. Learning experts will provide suggestions and opinions on the developed learning evaluation media materials.

d. Field Trial

6th semester students of the FITK study program at UIN Maulana Malik Ibrahim Malan, totaling 30 students.

3) Data Type

There are two types of developmental data: quantitative data and qualitative data. Quantitative data was obtained from survey and test results regarding student understanding. Quantitative data were collected through expert and practitioner surveys, student response surveys, and pre- and post-test result surveys. Qualitative data is in the form of descriptive quantitative data

4) Data Collection Timeline

The data collection process unfolded over a five-month period, following a structured sequence consistent with established Research and Development (R&D) procedures. During Month 1, the study commenced with a comprehensive needs analysis, which involved preliminary student surveys, interviews, and a systematic review of contemporary literature on learning evaluation and the Merdeka Curriculum. This phase aimed to identify existing

gaps in assessment resources and to delineate the conceptual foundation for the textbook to be developed.

Month 2 was dedicated to the design and initial development of the prototype. At this stage, the researchers formulated the conceptual framework of the textbook, organized its chapter structure, and drafted preliminary content, including evaluation instruments, example rubrics, and curriculum-aligned assessment procedures. This ensured that the prototype aligned with the competencies required in evaluation courses at the higher education level.

Month 3 focused on expert validation. The preliminary textbook was evaluated by a panel consisting of material experts, a design expert, and a practitioner with specialized experience in assessment implementation. The experts provided structured quantitative validation scores and qualitative recommendations regarding content accuracy, theoretical alignment, design clarity, and pedagogical suitability.

In Month 4, the textbook underwent refinement based on expert feedback and proceeded to field testing with a cohort of 30 sixth-semester students. During the field trial, students used the textbook over an extended period in real instructional settings. Observational data, student response questionnaires, and pretest–posttest measures were collected to evaluate the practicality and usability of the material.

Finally, Month 5 involved the integration of all data into a comprehensive evaluation. Quantitative data—including validation results, attractiveness scores, and learning gains—were analyzed using appropriate statistical techniques, while qualitative feedback was synthesized to inform the final revision of the product. This stage ensured that the textbook met the criteria of validity, practicality, and effectiveness prior to formal dissemination.

5) Data Collection Instruments

The study used four types of instruments:

a. Observation sheet (for expert and field observations).
b. Questionnaire (Likert scale) for expert validation and student responses.
c. Test (pretest–posttest) to measure improvement in students' understanding.
d. Documentation of development procedures and revisions.

6) Data Analysis Technique

a. Qualitative data analysis

Qualitative Data Analysis This analysis is a descriptive analysis based on the Miles and Hubermasn model. Data were obtained from observation and included reviews from material experts, design experts, and practitioners. The stages of data analysis include data processing, data reduction, data presentation, and conclusion drawing or validation. Below is a description of qualitative data analysis when developing textbook research for learning evaluation.

b. Quantitative data analysis

1) Validation analysis of learning evaluation textbooks

Textbooks related to the curriculum. The curriculum textbooks

were verified first before being tested. The researcher chose three categories of validation: content, design, and practitioner validation. Testers will check the set of questions according to their evaluation criteria. The validation data is available in the form of quantitative data.

2) Analyzed students' responses to the lesson evaluation book (curriculum). The student survey was used as a tool to measure the attractiveness of the lesson planning textbook.
3) Analysis of pretest and posttest results

The data obtained from the pre and post test evaluation results will be used to determine the degree of comparison between learning evaluation textbooks before and after publication. At the data analysis stage, the paired sample t-test formula is used (T-test) because the data is collected by determining two different measures from the same person. To avoid errors in the t test analysis, researchers used the SPSS 25.0 application.

RESULTS & DISCUSSION

A. Learning Evaluation Textbook Development Results

1. Description of the Learning Evaluation Textbook Based on the Merdeka Curriculum

The textbook developed by the researcher is an Merdeka Curriculum-based learning evaluation textbook that is able to improve the ability of good knowledge and skills in developing learning evaluations for PAI students. This Merdeka Curriculum-Based Learning Evaluation Textbook provides a comprehensive guide for teachers to understand, design, and implement evaluations that are in accordance with the philosophy and principles of the Merdeka Curriculum. With an evaluation approach that focuses more on student competence, character, and practical skill development, this book aims to improve the quality of learning and student success in achieving broader and holistic educational goals. Identity of Learning Evaluation Textbook.

Title	: Learning Evaluation in accordance with the Merdeka Curriculum (Concepts, Principles, and Procedures)
Material	: There are 5 chapters
Target	: Islamic Education Students
Supporting Software	: Microsoft Word and Adobe Photoshop
Book Size	: B5
Completeness of content	: Accompanied by an example form
Developer Name	: Drs. A. Zuhdi, MA Rasmuin, M.Pd.I Yuniar Setyo Marandy, S.Sn

a. Textbook Display

The following is the cover, table of contents, and excerpt of the contents of the Merdeka Curriculum-based learning evaluation textbook.

Figure 1. Cover of Learning Evaluation Book

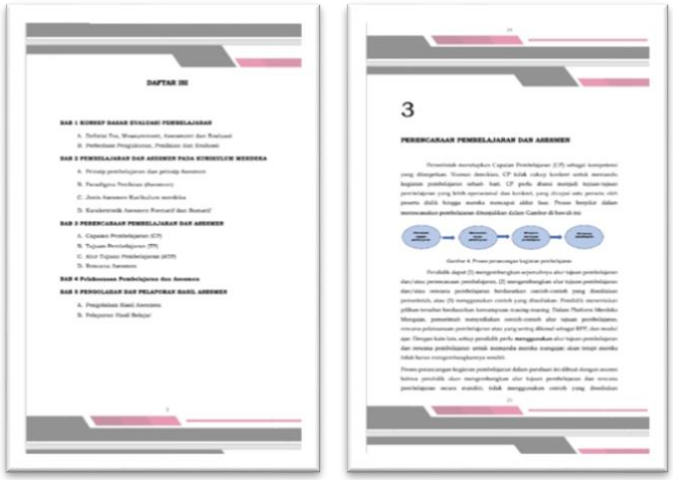

Figure 2. Table of Contents

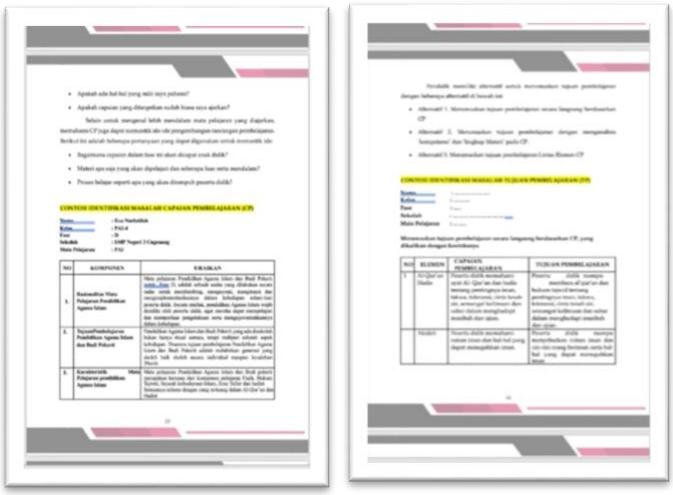

Figure 3. Book content display

The Merdeka Curriculum-based Learning Evaluation textbook is usually compiled to provide clear and easy-to-understand information for practitioners and students. The design and appearance of this book is very important to help readers understand the concepts and applications of learning evaluation in the context of the Merdeka Curriculum.

The book cover is the first element the reader will see, so the cover should reflect

the topic of the book and attract attention. Some elements that can be on the book cover are the Title of the Merdeka Curriculum-Based Learning Evaluation Book, Illustrations Images or designs that illustrate the concept of evaluation, education, or diversity for example, images showing students engaged in various learning activities (e.g., discussions, presentations, group work, projects), bright and fresh colors that reflect the spirit of an Merdeka Curriculum that is innovative and focuses on freedom and creativity.

The table of contents provides an overview of the structure of the book. Each chapter or main topic is clearly explained and organized so that readers can find the information they need easily. This book can be supplemented with appendices containing examples of assessment rubrics**,** evaluation question formats**,** or portfolio templates that can be used directly by teachers in learning evaluation.

The layout of the Merdeka Curriculum-based learning evaluation textbook is designed to provide readers with a thorough understanding of the principles and applications of evaluation in the Merdeka Curriculum. The clear, structured, and easy-to-understand design will help students implement evaluations that are effective, fair, and focused on developing student competence and character.

B. Learning Evaluation Textbook Validation Results

1. Material Expert Validation

Table 2. Material Expert Validation

Validator Type	N Items	Mean Score	Percentage (%)	Category
Material Expert	25	4.12	82.4%	Very Feasible
Design Expert	25	4.20	84.0%	Very Feasible
Practitioner Expert	25	4.20	84.0%	Very Feasible

Based on the data above, the results of the analysis of the material expert's assessment get a percentage of 82.4%, meaning that the learning evaluation textbook based on the Merdeka Curriculum is categorized as suitable for use as a textbook for students to support learning evaluation courses and is ready to be tested. Input from material experts during the validation process is to add more examples in the assessment and learning outcomes report of the Merdeka Curriculum. This is done so that students understand better with examples in developing assessments.

2. Design Expert Validation

Table 3. Design Expert Validation

Aspect	Score (Max)	Percentage	Interpretation
Layout organization	18/22	82%	Very Feasible
Visual clarity	17/20	85%	Very Feasible
Navigation & usability	20/24	83%	Very Feasible
Pedagogical design	14/16	88%	Very Feasible
Overall	–	84.0%	Very Feasible

The results of the analysis of the design expert's assessment obtained a percentage of 84%, meaning that the learning evaluation textbook is categorized as feasible to use as a book that is able to improve the ability to understand and skills in compiling learning evaluations for PAI students and is ready to be tested. Notes given by design experts include that good learning design must be based on a deep understanding of the psychology of learning, the context of students, and the application of relevant theory and practice to improve learning effectiveness. Overall, the book is fit for purpose.

3. Practitioner Expert Validation

Table 4. Practitioner Expert Validation

Aspect	Score (Max)	Percentage	Interpretation
Practicality	19/22	86%	Very Feasible
Clarity of instructions	17/20	85%	Very Feasible
Applicability in lectures	20/24	83%	Very Feasible
Alignment with student needs	14/16	88%	Very Feasible
Overall	–	84.0%	Very Feasible

The results of the analysis of the expert practitioner's assessment obtained a percentage of 84%, meaning that the coursebook is categorized as feasible to use as a companion book for learning evaluation lectures in terms of increasing the ability to understand and skills in linking assessment or evaluation of learning based on an Merdeka Curriculum. According to practitioners, no one has previously conducted research like this and coursebooks are feasible to implement in activities to improve the ability to understand and skills in linking Merdeka Curriculum-based learning evaluation. Evaluation is used to provide useful feedback for students and teachers. In the context of textbooks, evaluation often serves to measure student understanding of the material presented in the textbook. This feedback can be in the form of tests, quizzes or assignments that can help students understand concepts better and correct errors in their understanding.

C. Level of Attractiveness of the Textbook

The level of attractiveness of the textbook is determined by the percentage of the response questionnaire of 6th semester students as the main subject of the study. This questionnaire was filled in after the students were given treatment in the form of textbooks and learned to use this book for 3 months. The questionnaire was given to a sample of 30 students. Each student gets 10 questions related to the textbook. The calculation results in the student response questionnaire obtained a percentage of 85.3%.

D. The Effect of Learning Evaluation Textbooks Based on the Merdeka Curriculum to Improve Students' Understanding and Skills

To find out whether the textbook is able or not to improve students' understanding and skills, it is measured based on the results of the pretest posttest and the results of observations during the research process. Giving pretest questions is given to students before receiving coursebook treatment. This is done to determine the extent of students' ability to develop learning assessments. While the posttest questions are given after students study the textbook.

1. Descriptive and inferential data results

Descriptive data is used to determine the results of students' pretest and posttest scores simply after being accumulated. The following are the results of the assessment before (pretest) and after (posttest) giving textbooks.

The difference in the average pretest and posttest scores of these students simply shows that textbooks have an influence on learning evaluation skills. However, this cannot be a strong benchmark in other words, the results of the description data have not been able to provide conclusions about the results of observations made by researchers. Therefore, researchers utilize the SPSS 25.0 application to calculate mean, median, mode, standard deviation, variance, range, minimum maximum value, data normality test and hypothesis testing.

Furthermore, this pretest and posttest data is calculated using inferential data

analysis to determine the normal distribution of data and to see whether or not there is an effect of textbooks on research subjects in more detail. The analysis of pretest results is as follows:

Table 5. Analysis of Pretest and Posttest Results

		Mean	**N**	**Std. Deviation**	**Std. Error Mean**
Pair 1	pretest	72,00	30	2,729	,498
	posttes	88,67	30	4,823	,881

Table 6. Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Pretest & posttes	30	,650	,000

Table 7. Paired Samples Test

	Paired Differences					t	f	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1	pretest-posttes	-16,667	6,915	1,262	-19,249	-14,085	-13,201	29

a. Data normality test

This data analysis is used to find out whether the data obtained includes normal data or not. The normality test of the data uses the Kolmogorof Smirnow model because the number of samples used is 30. The rules of the normality test are "if the significance value> 0.05 then the data is normally distributed, and vice versa if the significance value <0.05 then the data is not normally distributed." The results of the data normality test were obtained:

Table 8. Test Of Normality

	Kolmogorov–Smirnov			**Shapiro–Wilk**		
Variable	Statistic	df	Sig.	Statistic	df	Sig.
Pretest	0.121	30	0.200*	0.967	30	0.435
Posttest	0.108	30	0.200*	0.973	30	0.582

b. Hypothesis Test

Based on the data above, it is known that the average (mean) on the pretest is 72. While the average on the posttest is 88. The data above shows that the average from pretest to posttest increased. The table above also shows the results of the standard deviation of the pretest of 2.729 and the standard deviation of the posttest of 4.823. The deviation distance on the pretest results is smaller than the deviation distance on the posttest. This means that before being given treatment in the form of textbooks, the average understanding and skill abilities are almost the same considering that the samples taken using purposive sampling technique by prioritizing students who are less able in the desired abilities.

From the above information, it can be concluded that the use of textbooks can improve the ability to understand and skills in preparing learning evaluations based on the Merdeka Curriculum for PAI students.

In the table above, the significance value is 0.000 <0.05. This result shows

that there is a significant effect between before being treated with textbooks and after. The conclusion obtained from the overall calculation is that this textbook can effectively improve comprehension and skills.

c. Effect Size

The effect size was also calculated to determine the magnitude of improvement. Using Cohen's d for repeated measures ($d = t / \sqrt{N}$), the intervention produced a very large effect size of $d = 2.57$, indicating that the curriculum-based evaluation textbook had a substantial impact on improving students' evaluation skills and conceptual understanding.

2. Observation Results of Improving Comprehension Ability and Skills Using Teaching Books

The findings of this study indicate that the Merdeka Curriculum–based evaluation textbook produced meaningful improvements in students' comprehension and assessment skills. Throughout the observation stages, students showed a clearer understanding of evaluation types and demonstrated increased ability to apply authentic and competency-based assessment principles. Before using the textbook, students tended to associate evaluation solely with written tests, but after the intervention, they were able to explain formative and summative functions and apply them in appropriate learning contexts. They also gained a more holistic view of evaluation through feedback, reflection, and continuous improvement, which aligns with the orientation of the Merdeka Curriculum.

These findings are consistent with recent research by Sokhanvar et al. (2021); Thorat (2025); Vlachopoulos and Makri (2024) shows that authentic assessment strengthens students' 21st-century competencies and supports deep learning an outcome parallel to the improvement observed in this study. Likewise, Papanastasiou et al. (2025) emphasize that authentic assessment facilitates sustainable learning and helps learners construct knowledge through reflective processes, which aligns with the shift in students' understanding found in this research. The improvement in students' evaluation skills also corresponds with the growing consensus that meaningful assessment practices require integrating feedback, contextualized tasks, and student-centered procedures, as highlighted by the findings of (Oroh & Ali, 2025).

The results of this study also align with recent developments in textbook effectiveness. Studies such as Li and Wang (2024) show that well-designed textbooks significantly improve students' academic performance and engagement. The increased comprehension and skills observed among students in this research support this view, demonstrating that structured and curriculum-aligned textbooks play an essential role in improving learning outcomes (Ham & Heinze, 2018). These findings also extend earlier studies by providing evidence in the specific context of the Merdeka Curriculum, a curricular framework that emphasizes authentic assessment and deeper learning.

This research provides a novel contribution by developing what appears to be one of the first evaluation textbooks explicitly aligned with the Merdeka Curriculum. While several recent studies have developed authentic assessment models and tools, such as Papanastasiou et al. (2025) and Vlachopoulos and

Makri (2024), none have focused on producing an integrated textbook that guides students through the philosophy, structure, and implementation of Merdeka Curriculum–based evaluation. The textbook created in this study includes competency-based rubrics, project and performance assessment examples, and model-based evaluation procedures, all of which are tailored to the characteristics of the Merdeka Curriculum. This contributes to filling the gap in instructional materials for pre-service teachers in Indonesia.

Despite these strengths, the study has certain limitations that should be acknowledged. The field testing was conducted only with one cohort of sixth-semester Islamic Education students, which limits the generalizability of the findings to other study programs or institutions. The number of expert validators was also limited, and most came from similar academic domains, potentially influencing the textbook evaluation results. Moreover, the implementation period lasted only eight weeks; thus, the long-term impact of the textbook on students' professional competence remains unknown. The data collected also relied primarily on questionnaires, observation notes, and cognitive tests, without the inclusion of classroom performance assessments that could provide deeper triangulation.

Nevertheless, the results offer meaningful implications for the improvement of evaluation instruction in higher education. The textbook developed in this study demonstrates the potential of curriculum-aligned materials in strengthening students' understanding of evaluation practices within the Merdeka Curriculum. Future research should consider involving a wider range of study programs, conducting larger-scale field testing, and integrating digital-based versions of the textbook. It would also be beneficial to expand the range of authentic assessment examples across multiple subjects to further enhance the applicability and relevance of the textbook.

CONCLUSION

The findings of this study demonstrate that the learning evaluation textbook developed for the Merdeka Curriculum is both valid and pedagogically effective. Expert validation results indicated high feasibility across material, design, and practitioner assessments, with validity scores ranging from 82.4% to 84%, confirming that the content, structure, and pedagogical alignment of the textbook meet established standards for higher education learning materials. Furthermore, field testing revealed substantial improvements in students' comprehension and practical evaluation skills, as reflected in the increase from pretest to posttest scores and a large effect size (Cohen's d), indicating strong learning gains attributable to the use of the textbook.

Student response data also showed a high level of product attractiveness (85.3%), suggesting that the textbook is perceived as accessible, relevant, and supportive of independent learning. These combined results affirm that the textbook effectively bridges theoretical concepts and practical application in learning evaluation, thereby strengthening students' competence as prospective educators.

Despite these positive outcomes, this study has several limitations. The field

test involved a relatively small sample from a single study program, which may limit the generalizability of the findings. In addition, the duration of implementation was restricted to one semester, preventing an examination of long-term retention or transfer of skills. The study also did not compare the textbook with alternative instructional materials, which could provide deeper insights into its relative effectiveness.

The implications of this research highlight the need for continued development of curriculum-aligned evaluation resources that integrate authentic assessment practices and competency-based approaches. Future research may consider larger and more diverse samples, multi-institutional trials, longitudinal evaluations, and the development of accompanying digital or interactive versions of the textbook to enhance accessibility and engagement. Expanding the product to include teacher training modules and integration with classroom assessment systems may further strengthen its impact in improving learning evaluation competencies.

ACKNOWLEDGEMENT

The authors would like to thank all those who provided support and guidance throughout this research, especially colleagues and supervisors for their valuable assistance.

REFERENCES

Aka, K. A. (2019). Integration Borg & Gall (1983) and Lee & Owen (2004) models as an alternative model of design-based research of interactive multimedia in elementary school. *Journal of Physics: Conference Series*, *1318*(1), 12022.

Aris, N. M., Ibrahim, N. H., & Halim, N. D. A. (2025). Design and Development Research (DDR) Approach in Designing Design Thinking Chemistry Module to Empower Students' Innovation Competencies. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *44*(1), 55 – 68. https://doi.org/10.37934/araset.44.1.5568

Black, S., & Allen, J. D. (2019). Part 11: Learning Strategies. *Reference Librarian*, *60*(4), 288 – 303. https://doi.org/10.1080/02763877.2019.1626317

Carpenter, S. K., Pan, S. C., & Butler, A. C. (2022). The science of effective learning with spacing and retrieval practice. *Nature Reviews Psychology*, *1*(9), 496 – 511. https://doi.org/10.1038/s44159-022-00089-1

Castro-Alonso, J. C., de Koning, B. B., Fiorella, L., & Paas, F. (2021). Five Strategies for Optimizing Instructional Materials: Instructor- and Learner-Managed Cognitive Load. *Educational Psychology Review*, *33*(4), 1379 – 1407. https://doi.org/10.1007/s10648-021-09606-9

Fajri, S., Ulaini, N., & Susantri, M. (2023). Implementasi Kurikulum Merdeka pada Pembelajaran Sejarah. *Kaganga: Jurnal Pendidikan Sejarah Dan Riset Sosial Humaniora*, *6*(2), 387–397.

Febrianto, P. T., Nursalim, M., Bachri, B. S., Mustaji, & Mas'udah, S. (2025). Implementation, impact and strategies for facing challenges in the independent learning curriculum in elementary schools. *Multidisciplinary Science Journal*, *7*(7).

https://doi.org/10.31893/multiscience.2025414

Febrianto, R., & Puspitaningsih, F. (2020). Pengembangan buku ajar evaluasi pembelajaran. *Education Journal: Journal Educational Research and Development*, *4*(1), 1–18.

Fei, J., Zhao, J., Sun, S., & Liu, Y. (2018). Active learning methods with deep gaussian processes. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11303 LNCS*, 473 – 483. https://doi.org/10.1007/978-3-030-04182-3_41

Fiorella, L., & Mayer, R. E. (2015). Learning as a generative activity: Eight learning strategies that promote understanding. In *Learning as a Generative Activity: Eight Learning Strategies that Promote Understanding*. Cambridge University Press. https://doi.org/10.1017/CBO9781107707085

Goldman, S. R., & Pellegrino, J. W. (2015). Research on Learning and Instruction: Implications for Curriculum, Instruction, and Assessment. *Policy Insights from the Behavioral and Brain Sciences*, *2*(1), 33 – 41. https://doi.org/10.1177/2372732215601866

Han, Y. (2024). Research on the Influence of AI Technology on University Learning Effect and Optimization Strategy. *Learning and Analytics in Intelligent Systems*, *41*, 208 – 218. https://doi.org/10.1007/978-3-031-69457-8_20

Hima, L. R., Murniasih, T. R., Palayukan, H., & Shodikin, A. (2025). Differentiated Learning Strategies: Analysis of Merdeka Curriculum in Senior High Schools Mathematics Learning. In I. N. & S. Y. (Eds.), *AIP Conference Proceedings* (Vol. 3142, Issue 1). American Institute of Physics. https://doi.org/10.1063/5.0262767

Japa, I. G. N. (2012). Pengembangan Buku Ajar Berpendekatan Pendidikan Matematika Realistik dan Pemecahan Masalah Terbuka. *Jurnal Pendidikan Dan Pengajaran*, *45*(2).

Jiménez-Macías, A., Muñoz-Merino, P. J., & Kloos, C. D. (2024). Student Clustering Through KMeans to Enhance Teaching. *16th Congreso de Tecnologia, Aprendizaje y Ensenanza de La Electronica, TAEE 2024*. https://doi.org/10.1109/TAEE59541.2024.10604956

Marlina, L., & Sholehun, S. (2021). Analisis faktor-faktor yang mempengaruhi hasil belajar bahasa indonesia pada siswa kelas iv sd muhammadiyah majaran kabupaten sorong. *FRASA: Jurnal Keilmuan, Bahasa, Sastra, Dan Pengajarannya*, *2*(1), 66–74.

May, M.-D., Lin, H.-Y., Tseng, Y.-T., & Sung, C.-M. (2020). Application of PBL in ERP textbook design and technical university students learning effectiveness study. In C. A., A. G.R., & R. T. (Eds.), *IEEE Global Engineering Education Conference, EDUCON* (Vols. 2020-April, pp. 878 – 884). IEEE Computer Society. https://doi.org/10.1109/EDUCON45650.2020.9125177

Nintarat, L., Narongkul, S., & Akatimagool, S. (2023). A Design and Development of Learning Innovations for Training Courses on Smart Greenhouse Systems. *8th International STEM Education Conference, ISTEM-Ed 2023 - Proceedings*. https://doi.org/10.1109/iSTEM-Ed59413.2023.10305746

Prastowo, A. (2011). *Panduan kreatif membuat bahan ajar inovatif*. Yogyakarta: DIVA press.

Ramli, N., Al Hafis Hashim, M. E., & Mustafa, W. A. (2024). An Interactive AR Textbook Prototype Conceptual Framework Assesses Elementary School Students' Learning Experiences and Emotions Using Design-Based Research (DBR). In M. W.A.W., I. S.Z.S., K. N.F., A. K.N.F.K., & Y. N. (Eds.), *AIP Conference Proceedings* (Vol. 3135, Issue 1). American Institute of Physics. https://doi.org/10.1063/5.0212677

Ruliah, Muslim, S., Syahrial, Z., & Pratiwi, A. S. (2020). Development of instructional design based on computer assisted instruction models for database system course in

information technology colleges. *Universal Journal of Educational Research*, *8*(9 A), 21 – 30. https://doi.org/10.13189/ujer.2020.082004

Salmia, Nursalam, & Bancong, H. (2024). Effectiveness of Local Wisdom-Based Merdeka Curriculum Teaching Modules in Improving Learning Outcomes Indonesia. *Journal of Ecohumanism*, *3*(6), 1719 – 1726. https://doi.org/10.62754/joe.v3i6.4131

Sukriadi, S., Kartono, K., & Wiyanto, W. (2015). Analisis Hasil Penilaian Diagnostik Kemampuan Berpikir Kritis Matematis Siswa Dalam Pembelajaran Pmri Berdasarkan Tingkat Kecerdasan Emosional. *Unnes Journal of Mathematics Education Research*, *4*(2).

Surbakti, R. S. (2017). *PENGEMBANGAN BAHAN AJAR MENULIS TEKS EKSPLANASI BERBASIS KONTEKSTUAL KELAS VII SMP NEGERI 1 PEMATANGSIANTAR TAHUN PEMBELAJARAN 2016/2017*. UNIMED.

Wang, J. (2025). Research on Large Model Resource Optimization Strategy Through Distributed Management. *Proceedings - 2025 5th Asia-Pacific Conference on Communications Technology and Computer Science, ACCTCS 2025*, 535 – 541. https://doi.org/10.1109/ACCTCS66275.2025.00098

Zaini, M. (2006). Pengembangan Kurikulum: Konsep. *Implementasi Evaluasi Dan Inovasi (Yogyakarta*.

Zhao, Y.-T.-Y., Jia, Z.-Y., Tang, Y., Xiong, J. J., & Zhang, Y.-C. (2018). Quantitative learning strategies based on word networks. *Physica A: Statistical Mechanics and Its Applications*, *491*, 898 – 911. https://doi.org/10.1016/j.physa.2017.09.097