International Journal of Education, Vocational and Social Science
https://e-journal.citakonsulindo.or.id/index.php/IJEVSS
E-ISSN : 2963 4954
IJEVSS

Development Of A School Mathematics Exam Test Instrument For Grade IX Of Selayar Islands State Middle School

Nur Kumala Sari[1]
[1]Universitas Negeri Makassar, Indonesia
Syahrul[2]
[2]Universitas Negeri Makassar, Indonesia
Baso Intang Sappaile[3]
[3]Universitas Negeri Makassar, Indonesia

Corresponding Author: fildzahfatiyah@gmail.com ; syahrul@unm.ac.id; baso.sappaile@unm.ac.id

Abstract

This study aims to develop a high-quality school mathematics examination test instrument for Grade IX students at Selayar Islands State Middle School. The research employs a research and development (R&D) method based on the Borg and Gall model, involving systematic stages from initial design, validation, testing, to final product revision. The results indicate that the developed instrument meets the criteria of a good assessment tool, with 20 out of 30 items declared valid and the model achieving goodness-of-fit standards, demonstrating accurate construct measurement. The instrument also shows high reliability, with coefficient α and ω values of 0.839 and an average inter-item correlation of 0.208. In terms of difficulty level, most items fall into the moderate category, ensuring proportional measurement of students' abilities, with only one item classified as difficult. Furthermore, the discrimination power of most items is categorized as good to very good, although a few require revision. Overall, the instrument is valid, reliable, and effective, making it suitable for objectively and consistently evaluating students' mathematics learning outcomes.

Keywords: Mathematics Test Instrument, School Exam, Grade IX, Test Development

INTRODUCTION

Education plays a fundamental role in shaping high-quality human resources who are

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

capable of adapting to the rapid development of science, technology, and global challenges. In the context of secondary education, mathematics is recognized as a core subject that contributes significantly to the development of students' logical, analytical, and systematic thinking skills. As emphasized in the thesis, mathematics learning at the junior high school level is not merely oriented toward mastering concepts and computational skills, but also toward fostering higher-order thinking skills, including reasoning, critical thinking, and problem-solving abilities. These competencies are essential in preparing students to face real-life problems and to participate effectively in a knowledge-based society. Therefore, the quality of mathematics education must be supported by an effective and comprehensive assessment system that accurately reflects students' learning achievements(Elizondo-Montemayor, 2004).

Assessment, particularly in the form of school examinations, serves as a crucial component in the educational process. It functions not only as a tool to measure students' learning outcomes but also as an instrument to evaluate the effectiveness of teaching and learning practices. School examinations, as a form of summative assessment, are designed to measure the extent to which students have achieved the expected competencies at the end of a learning period. Consequently, the quality of the test instruments used in such examinations becomes a determining factor in ensuring the validity and credibility of the assessment results. A well-developed test instrument should meet psychometric criteria, including validity, reliability, appropriate difficulty level, and good discrimination power, so that it can provide accurate and meaningful information about students' abilities(Blumberg, 2005).

However, the reality in many schools, particularly in remote areas such as the Selayar Islands, indicates that the quality of school examination instruments is still relatively low. Based on the findings presented in the thesis, many mathematics test items are developed using conventional approaches that predominantly measure low-level cognitive skills. Moreover, the process of test construction often lacks empirical analysis, such as validity testing, reliability analysis, item difficulty, and discrimination indices. This condition leads to the use of test instruments that are not fully capable of representing students' actual competencies. As a result, the outcomes of school examinations may not accurately reflect students' true learning achievements, thereby reducing the effectiveness of assessment as a decision-making tool in education(Jaenudin & Murwaningsih, 2017).

Another significant issue highlighted in the study is the limited professional competence of teachers in developing and analyzing assessment instruments. Many teachers have not received sufficient training in constructing high-quality test items based on established

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

educational standards and psychometric principles. In addition, the lack of continuous guidance and support from educational authorities contributes to teachers' reliance on outdated question banks without proper validation processes. This situation is further exacerbated by geographical challenges in the Selayar Islands, where limited access to professional development opportunities and educational resources hinders the improvement of assessment practices(van Der Vleuten & Schuwirth, 2019).

From a theoretical perspective, a good test instrument must adhere to the principles of educational evaluation and measurement. Classical Test Theory (CTT) emphasizes the importance of item analysis to ensure that each test item contributes effectively to measuring students' abilities. In addition, modern educational assessment requires the integration of higher-order thinking skills (HOTS) into test items to align with current curriculum demands. Therefore, the development of test instruments should not only focus on statistical quality but also on their relevance to curriculum objectives and students' contextual characteristics(Andrijati et al., 2025).

Previous studies have attempted to develop valid and reliable mathematics test instruments; however, most of these studies are conducted in urban settings with relatively adequate educational facilities. As a result, there is still a gap in research focusing on the development of school examination instruments in geographically unique areas such as island regions. Furthermore, the development of summative assessment instruments specifically designed for school examinations at the junior high school level remains limited. This gap highlights the need for research that integrates theoretical frameworks of educational evaluation with practical development models to produce contextually relevant and high-quality test instruments(Yendra et al., 2023).

Based on these considerations, it is necessary to develop a mathematics school examination test instrument for ninth-grade students that is scientifically designed, systematically developed, and contextually relevant to the characteristics of students in the Selayar Islands. Such development should integrate the principles of Classical Test Theory with a research and development approach to produce an instrument that is valid, reliable, and practical for use in real educational settings. Ultimately, the development of a high-quality test instrument is expected to improve the accuracy of learning outcome measurements, enhance the quality of assessment practices, and contribute to the overall improvement of mathematics education, particularly in remote and underserved regions.

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

The Concept of Educational Evaluation and Learning Assessment

Educational evaluation and learning assessment are fundamental components in ensuring the effectiveness and quality of the teaching and learning process. In the context of modern education, evaluation is not merely an activity of measuring students' achievement, but rather a systematic process of collecting, analyzing, and interpreting data to determine the value and success of educational programs. As described in the thesis, evaluation involves making judgments and decisions based on empirical evidence and predetermined criteria, thereby serving as a reflective mechanism for continuous improvement in education. Assessment, on the other hand, is more specifically focused on gathering information related to students' learning outcomes through various instruments such as tests, observations, and assignments. It provides the operational foundation for evaluation by supplying the necessary data for decision-making processes(Taradi et al., 2005).

The relationship between evaluation, assessment, testing, and measurement is hierarchical in nature. Measurement refers to the process of assigning numerical values to students' performance, while testing is a tool used to obtain such measurements. Assessment encompasses broader methods of collecting information about learning, and evaluation uses this information to make informed judgments about educational effectiveness. This integrated framework ensures that the educational process is aligned with learning objectives and that decisions are evidence-based. Furthermore, effective assessment must align with intended learning outcomes, ensuring that what is taught, what is assessed, and what is expected from students are coherent and consistent(Septiana & Hermanto, 2024).

In contemporary educational practices, assessment serves both formative and summative purposes. Formative assessment provides ongoing feedback to improve learning during the instructional process, while summative assessment evaluates students' achievement at the end of a learning period. The thesis emphasizes that both forms are essential and must be balanced to avoid an overemphasis on final results at the expense of meaningful learning processes. Additionally, assessment must adhere to key principles such as validity, reliability, fairness, transparency, and usefulness. Validity ensures that the instrument measures what it is intended to measure, while reliability guarantees consistency in results across different contexts and times(Masek & Yamin, 2010).

In the Indonesian educational system, these principles are embedded within national assessment standards, which require alignment between assessment instruments and curriculum objectives. Assessment is also expected to support accountability and quality assurance in

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

education. Therefore, a well-designed assessment system not only measures learning outcomes but also contributes to improving teaching practices and educational policies. By integrating these concepts, educational evaluation becomes a powerful tool for enhancing the overall quality of education and ensuring that students achieve the competencies required in the 21st century.

School Examinations as a Form of Summative Assessment

School examinations represent one of the most important forms of summative assessment in the education system. They are designed to measure students' achievement of learning objectives at the end of a specific instructional period and serve as a basis for important academic decisions, such as graduation and certification. According to the thesis, school examinations function as an accountability mechanism for schools to demonstrate the effectiveness of their educational processes and the extent to which students have mastered the required competencies. Therefore, the quality of the examination instruments directly influences the credibility and accuracy of the assessment results(Alghasham, 2012).

A high-quality school examination must be developed based on clear specifications derived from the curriculum. This includes defining learning objectives, identifying key competencies, and ensuring appropriate representation of content areas. The use of a test blueprint or table of specifications is essential in this process, as it ensures that the test items are distributed proportionally across different topics and cognitive levels. This approach helps maintain content validity and ensures that the test accurately reflects the intended learning outcomes.

Moreover, school examinations should not be limited to measuring lower-order cognitive skills such as memorization and basic comprehension. Instead, they should incorporate higher-order thinking skills (HOTS), including analysis, evaluation, and problem-solving. The thesis highlights that many existing school examination items still focus on procedural knowledge, thereby failing to capture students' deeper understanding and critical thinking abilities. To address this issue, test developers must design items that require students to apply concepts in real-life contexts and engage in complex reasoning processes(SAPUTRA et al., 2023).

Another critical aspect of school examinations is fairness and objectivity. Test items must be free from bias and accessible to all students, regardless of their background. Clear instructions, appropriate language use, and well-constructed answer options are necessary to ensure that the test measures students' knowledge rather than their ability to interpret ambiguous questions. In addition, standardized administration procedures are required to maintain consistency across different schools and testing conditions(Li et al., 2023).

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

The role of school examinations extends beyond individual assessment to system-level evaluation. The results can be used to identify strengths and weaknesses in the curriculum, inform instructional improvements, and guide policy decisions. However, the effectiveness of school examinations depends heavily on the quality of the test instruments used. Poorly constructed tests can lead to inaccurate conclusions about students' abilities and undermine the validity of the entire assessment system. Therefore, the development of high-quality school examination instruments is a critical priority in improving educational outcomes.

Development of Test Instruments Based on Classical Test Theory

The development of high-quality test instruments requires a strong theoretical foundation, and one of the most widely used frameworks in educational measurement is Classical Test Theory (CTT). According to the thesis, CTT provides a practical and systematic approach for analyzing the quality of test items and ensuring that they function effectively as measurement tools. The central concept of CTT is that an observed test score consists of a true score and an error component. Therefore, the goal of test development is to minimize measurement error and maximize the accuracy of the scores(Yuliati[1] et al., 2024).

One of the key components of CTT is item analysis, which involves examining various statistical indicators to evaluate the quality of test items. These indicators include item validity, reliability, difficulty level, discrimination index, and distractor effectiveness. Item validity refers to the extent to which a test item measures the intended construct, while reliability indicates the consistency of the test results. High reliability suggests that the test produces stable and dependable measurements.

The difficulty level of an item indicates how easy or difficult it is for students, typically expressed as the proportion of students who answer the item correctly. A well-balanced test should include items with varying levels of difficulty, but most items should fall within the moderate range to effectively differentiate between students of different ability levels. The discrimination index measures the ability of an item to distinguish between high-performing and low-performing students. Items with high discrimination power are considered more effective because they contribute to the overall accuracy of the test(Aufa et al., 2020).

Another important aspect of item analysis is distractor effectiveness. In multiple-choice tests, distractors (incorrect answer options) must be plausible and attractive to students who do not know the correct answer. Effective distractors help ensure that the test measures students' knowledge rather than guessing ability. The thesis emphasizes that all these indicators must be

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

carefully analyzed and used as a basis for revising and improving test items(Dewi & Mercuriani, 2024).

The process of developing test instruments based on CTT typically involves several stages, including needs analysis, test design, item construction, expert validation, field testing, and revision. This systematic approach ensures that the final product meets both theoretical and empirical standards of quality. By integrating CTT principles into the development process, educators can produce test instruments that are not only statistically sound but also practical and applicable in real classroom settings. Ultimately, the use of CTT contributes to the development of fair, accurate, and meaningful assessment tools that support effective learning and teaching(Martaningsih et al., 2022).

Challenges in Developing Mathematics Test Instruments in Island Regions

Developing high-quality mathematics test instruments in geographically isolated regions, such as the Selayar Islands, presents unique challenges that require special attention. The thesis highlights several factors that affect the quality of assessment practices in these areas, including limited resources, geographical constraints, and variations in teachers' professional competencies. These challenges have significant implications for the development and implementation of school examination instruments.

One of the primary challenges is the limited access to professional development opportunities for teachers. Due to geographical isolation, teachers in island regions often have fewer opportunities to participate in training programs related to assessment and test development. As a result, many teachers lack the necessary skills and knowledge to construct high-quality test items and conduct proper item analysis. This situation leads to the use of conventional test items that do not meet psychometric standards(Dewanti, 2024).

Geographical factors also play a crucial role in shaping educational practices. The dispersed nature of island regions makes it difficult to organize training sessions, conduct supervision, and facilitate collaboration among schools. This lack of interaction limits the exchange of best practices and hinders the development of standardized assessment instruments. Additionally, infrastructure limitations, such as inadequate access to technology and learning resources, further complicate the implementation of modern assessment approaches.

Another challenge is the contextual relevance of test instruments. Many existing test items are developed based on general contexts that may not reflect the realities of students in island regions. As a result, students may struggle to understand or relate to the questions, which can

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

affect their performance. The thesis emphasizes the importance of developing context-based assessment instruments that consider students' backgrounds, experiences, and local environments. This approach ensures that the assessment is not only valid but also meaningful for students(Khoiriyah & Roberts, 2025).

Furthermore, there is often a lack of standardized instruments that have been empirically validated for use in these regions. Without such instruments, schools rely on locally developed tests that may vary in quality and consistency. This situation creates disparities in assessment practices and makes it difficult to compare students' performance across different schools. To address this issue, there is a need for systematic research and development efforts that produce validated and reliable test instruments tailored to the specific context of island regions(Hayati et al., 2020).

In conclusion, addressing these challenges requires a comprehensive approach that combines theoretical knowledge, practical training, and contextual adaptation. By developing test instruments that are scientifically grounded and contextually relevant, educators can improve the quality of assessment and support better learning outcomes for students in remote areas. The thesis demonstrates that with a structured development model and the application of Classical Test Theory, it is possible to overcome these challenges and produce high-quality mathematics test instruments that meet educational standards and local needs.

DISCUSSION

The Empirical Condition of the Implementation of Mathematics School Examinations

The empirical condition of the implementation of mathematics school examinations for ninth-grade students in junior high schools in the Selayar Islands reveals several fundamental issues related to the quality of the test instruments currently used. Based on the findings presented in Chapter IV, the existing examination instruments tend to emphasize lower-order cognitive skills, primarily focusing on recall and basic understanding rather than higher-order thinking skills such as analysis and problem-solving. The test items are often constructed without a comprehensive blueprint or test specification table that ensures alignment between learning objectives, indicators, and assessment items. As a result, there is a mismatch between the competencies expected in the curriculum and what is actually measured in the examinations(Damayanti & Yohandri, 2022).

Furthermore, the development of these instruments rarely involves systematic validation processes, either through expert judgment or empirical testing. Many of the items used in school

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

examinations have not undergone proper item analysis, including validity testing, reliability estimation, difficulty index calculation, and discrimination power analysis. This lack of rigorous evaluation results in test instruments that may not accurately measure students' true abilities. Additionally, teachers often rely on previously used questions or question banks without revising them based on current curriculum demands or student characteristics. The absence of data-driven decision-making in constructing test items contributes to the persistence of low-quality instruments. These conditions are exacerbated by limited professional training opportunities for teachers in the field of educational assessment and evaluation. Consequently, the implementation of school examinations in mathematics in the Selayar Islands has not fully met the standards of a high-quality assessment system, highlighting the urgent need for systematic instrument development.

The Development Process of a Valid and Reliable Test Instrument Based on Borg and Gall Model

The development of the mathematics school examination test instrument in this study follows the Borg and Gall research and development model, which provides a systematic and iterative framework for producing high-quality educational products. The process begins with a needs analysis stage, where the researcher identifies the gaps and weaknesses in the existing examination instruments. This stage involves reviewing current test items, analyzing curriculum requirements, and gathering input from teachers and stakeholders. Following this, the planning stage is conducted by designing a test blueprint that outlines the competencies to be measured, the distribution of items across content areas, and the cognitive levels targeted in each item. The development stage involves constructing a pool of test items based on the blueprint, ensuring that each item reflects the intended indicators and adheres to principles of good item writing. Subsequently, the initial product undergoes expert validation involving subject matter experts, evaluation experts, and language experts who assess the content validity, construct validity, and linguistic clarity of the items. Revisions are made based on the feedback obtained from these experts. The next stage involves field testing, where the instrument is administered to a sample of ninth-grade students in selected schools in the Selayar Islands. Data collected from this stage are analyzed using both classical test theory and confirmatory factor analysis (CFA) to evaluate the psychometric properties of the instrument. The iterative nature of the Borg and Gall model allows for continuous refinement of the instrument based on empirical evidence. The final stage involves producing a validated and reliable test instrument that is ready for practical implementation. This systematic development process ensures that the resulting instrument is

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

not only theoretically sound but also empirically tested and contextually relevant(Ramadani et al., 2024).

Results of Validity, Reliability, Item Difficulty, Discrimination Index, and Distractor Effectiveness

The results of the psychometric analysis of the developed test instrument indicate that the majority of the items meet the established criteria for a high-quality assessment tool. Content validity analysis conducted through expert judgment shows that the items are highly relevant to the curriculum and accurately represent the intended learning indicators. Construct validity, examined confirmatory factor analysis (CFA), confirms that the instrument aligns with the four main mathematical domains: numbers, algebra, geometry, and statistics. Most items demonstrate satisfactory factor loadings, indicating that they effectively measure the intended constructs. Empirical validity testing further supports these findings, with a significant proportion of items showing acceptable correlation coefficients with the total test score. In terms of reliability, the instrument achieves a high reliability coefficient, suggesting consistent measurement across items. The analysis of item difficulty reveals a balanced distribution, with items categorized as easy, moderate, and difficult, ensuring that the test can differentiate among students with varying ability levels. The discrimination index analysis shows that most items have good to very good discrimination power, meaning they can effectively distinguish between high-performing and low-performing students. Additionally, the effectiveness of distractors in multiple-choice items is generally satisfactory, with most distractors functioning properly by attracting responses from students with lower ability levels. However, a few items require revision due to non-functioning distractors or suboptimal discrimination indices. Overall, the results indicate that the developed instrument has strong psychometric properties and is suitable for use in measuring students' mathematical competencies(Sari et al., 2025).

Feasibility and Practicality of the Developed Test Instrument

The feasibility and practicality of the developed mathematics test instrument are evaluated based on expert assessments, teacher feedback, and field trial results. Expert validation results indicate that the instrument meets the criteria for content, construction, and language appropriateness. Subject matter experts confirm that the items accurately reflect mathematical concepts and are aligned with the curriculum, while evaluation experts highlight the sound structure and adherence to assessment principles. Language experts note that the items are clearly worded and easily understood by students, minimizing ambiguity and

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

misinterpretation(Pradipta et al., 2025). Teacher feedback further supports the feasibility of the instrument, as teachers find it relevant to classroom instruction and suitable for assessing students' learning outcomes. They also appreciate the inclusion of various cognitive levels, which allows for a more comprehensive evaluation of students' abilities. The practicality of the instrument is demonstrated through field trials, where the test can be administered efficiently within the allocated time without causing significant difficulties for students. Students report that the instructions and questions are clear, and the level of difficulty is appropriate. Moreover, the scoring process is straightforward, enabling teachers to efficiently analyze student performance. These findings suggest that the developed instrument is not only theoretically valid but also practically applicable in real classroom settings. The positive responses from both experts and practitioners indicate that the instrument has high usability and acceptability(Sari et al., 2025).

The Final Form of the Mathematics School Examination Test Instrument

The final form of the developed mathematics school examination test instrument represents a refined and validated product that meets the criteria of validity, reliability, and contextual relevance. The instrument consists of 30 multiple-choice items covering four main content areas: numbers, algebra, geometry, and statistics and probability. Each item is designed based on a detailed blueprint that ensures alignment with curriculum standards and learning objectives. The final selection of items is based on rigorous analysis, including expert validation and empirical testing, ensuring that only high-quality items are retained. The instrument incorporates a range of cognitive levels, from basic understanding to higher-order thinking skills, enabling a comprehensive assessment of students' mathematical abilities. Additionally, the contextualization of items reflects the characteristics of the Selayar Islands, making the test more relevant and meaningful for students. The language used in the instrument is clear, concise, and appropriate for the students' level of comprehension. The scoring system is standardized, allowing for consistent and objective evaluation of student performance. Overall, the final product demonstrates a significant improvement compared to the initial condition of test instruments used in schools. It provides a reliable and valid tool for measuring students' learning outcomes and can serve as a model for developing similar instruments in other regions. The successful development of this instrument highlights the importance of a systematic and research-based approach in improving the quality of educational assessment(Ssemugenyi, 2023).

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

CONCLUSION

Based on the results of the research and development process, the Grade IX Mathematics school examination instrument developed for SMP Negeri Kepulauan Selayar meets the standards of a high-quality assessment tool. Developed systematically through the Borg and Gall stages up to the final product revision, the instrument is considered appropriate for evaluating students' learning outcomes. Of the 30 items, 20 were found to be valid, and the model met the goodness-of-fit criteria, indicating accurate construct measurement. The instrument also demonstrated high reliability, with coefficient α and ω values of 0.839 and an average inter-item correlation of 0.208, categorized as good. The level of difficulty was predominantly moderate, allowing for proportional measurement of student ability, with only one difficult item. Additionally, most items showed good to very good discrimination power, although a few require revision. Overall, the instrument is valid, reliable, well-balanced in difficulty, and effective in distinguishing student performance, making it suitable as an objective, consistent, and effective tool for educational evaluation and decision-making.

FUTURE DIRECTIONS

Future directions for this study include expanding the development and validation of the Mathematics examination instrument by involving larger and more diverse student populations across different regions to enhance its generalizability. Further research is also recommended to refine the remaining items with lower validity and discrimination indices, as well as to integrate technology-based testing formats such as computer-based or adaptive assessments to improve efficiency and accuracy. In addition, longitudinal studies could be conducted to examine the stability of the instrument over time and its impact on students' learning outcomes. Incorporating higher-order thinking skills (HOTS) more comprehensively and aligning the instrument with evolving curriculum standards will also be important to ensure its continued relevance and effectiveness in supporting evidence-based educational decision-making.

REFERENCES

Alghasham, A. A. (2012). Effect of students' learning styles on classroom performance in problem-based learning. Medical Teacher, 34(sup1), S14–S19.

Andrijati, N., Yulianto, S., & Sukasih, S. (2025). Problem Based Learning in Developing Mathematical Reasoning Assessment Model for Sustainable Development. The 8th International Conference on Education Innovation (ICEI 2024), 637–650.

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

Aufa, M. N., Iriani, R., Saadi, P., Hasbie, M., Fitri, M. A., & Yunita, A. (2020). Module Development with Problem Based Learning (PBL) Model Based on Environmental Wetland to Increase Students' Learning Outcomes. JKPK (Jurnal Kimia Dan Pendidikan Kimia), 5(2), 201–210.

Blumberg, P. (2005). Assessing students during the problem-based learning (PBL) process. Journal of the International Association of Medical Science Educators, 15(2), 92–99.

Damayanti, I. R., & Yohandri, Y. (2022). E-Book Development Effectiveness Problem Based Learning with Quizing in Physics Learning. Jurnal Penelitian Pendidikan IPA, 8(6), 3044–3049.

Dewanti, S. S. (2024). Problem-Based Learning with Manipulative Media: Enhancing Elementary Students' Conceptual Understanding in Area and Volume. Jurnal Ilmiah Sekolah Dasar, 8(3), 487–499.

Dewi, R., & Mercuriani, I. S. (2024). Development of e-module in problem-based learning (pbl) on biotechnology material to increase learning independence and understanding of concepts for class xii sma students. Jurnal Penelitian Pendidikan IPA, 10(9), 6552–6563.

Elizondo-Montemayor, L. L. (2004). Formative and summative assessment of the problem-based learning tutorial session using a criterion-referenced system. Journal of the International Association of Medical Science Educators, 14(1), 8–14.

Hayati, U., Johar, R., & Marwan. (2020). The integration of local context through problem-based learning (PBL) to improve junior high school students problem-solving skills. Journal of Physics: Conference Series, 1460(1), 12036.

Jaenudin, A., & Murwaningsih, T. (2017). The effectiveness of the E-module of economics learning on problem-based learning used to improve students' learning outcomes. International Conference on Teacher Training and Education 2017 (ICTTE 2017), 290–296.

Khoiriyah, U., & Roberts, C. (2025). Investigating the role of self-assessment in enhancing self-regulated learning amongst medical students in problem-based learning. BMC Medical Education, 25(1), 780.

Li, A., Mellon, M., Keuhl, A., & Sibbald, M. (2023). Measuring group function in problem-based learning: development of a reflection tool. BMC Medical Education, 23(1), 745.

Martaningsih, S. T., Maryani, I., Prasetya, D. S., Purwanti, S., Sayekti, I. C., Aziz, N. A. A., & Siwayanan, P. (2022). STEM problem-based learning module: a solution to overcome elementary students' poor problem-solving skills. Pegem Journal of Education and Instruction, 12(4), 340–348.

Masek, A. Bin, & Yamin, S. Bin. (2010). Problem Based Learning: A review of the monitoring and

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

assessment model. 2010 2nd International Congress on Engineering Education, 171–175.

Pradipta, I. K. R., Agung, A. A. G., & Ambara, D. P. (2025). E-Book Interactive Problem Based Learning of the History of the Indonesian Nation's Struggle Material in Science Learning to Improve the Learning Outcomes of Grade V Elementary School Students: Bahasa Indonesia. MIMBAR PGSD Undiksha, 13(1), 12–22.

Ramadani, F., Desyandri, D., Sukma, E., & Erita, Y. (2024). Development of Interactive Multimedia Indonesian Language Teaching Material Based on Problem-Based Learning (PBL) in Elementary Schools. AL-ISHLAH: Jurnal Pendidikan, 16(2), 2257–2275.

SAPUTRA, B. H., PARNO, P., YULIATI, L., & Marlina, A. (2023). Building student's critical thinking skill with stem integrated problem based learning model through hybrid learning assisted by formative assessment on heat and temperature topic in the covid-19 pandemic era. Sciendo Conference Proceedings, 319, 90–97.

Sari, N., Armis, A., & Yuanita, P. (2025). Implementation of Problem Based Learning to Improve Students' Mathematics Learning Outcomes at UPT SMP Negeri 10 Tapung. Journal of Education and Learning Research, 3(1), 29–36.

Septiana, N., & Hermanto, F. Y. (2024). Construct validity of non-test instrument for problem-based learning method. Journal of Office Administration: Education and Practice, 4(1), 1–9.

Ssemugenyi, F. (2023). Teaching and learning methods compared: A pedagogical evaluation of problem-based learning (PBL) and lecture methods in developing learners' cognitive abilities. Cogent Education, 10(1), 2187943.

Taradi, S. K., Taradi, M., Radic, K., & Pokrajac, N. (2005). Blending problem-based learning with Web technology positively impacts student learning outcomes in acid-base physiology. Advances in Physiology Education, 29(1), 35–39.

van Der Vleuten, C. P. M., & Schuwirth, L. W. T. (2019). Assessment in the context of problem-based learning. Advances in Health Sciences Education, 24(5), 903–914.

Yendra, N., Fauzan, A., & Junedi, B. (2023). Development of problem based learning (PBL) learning tools to improve mathematical problem solving ability of class VIII SMP/MTs students. AIP Conference Proceedings, 2698(1), 60011.

Yuliati1, M. K. P., Jatmiko, B., & Ali, M. (2024). The Effect of STEM Approach with Assessment Formative in PBL on the Problem-Solving Ability of. Proceedings of the International Conference on Mathematics and Science Education (ICoMSE 2023), 118.

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License