Journal of Natural Science and Integration
P-ISSN: 2620-4967|E-ISSN: 2620-5092
Vol. 9, No. 1, April 2026, pp 84-96

Available online at:
http://ejournal.uin-suska.ac.id/index.php/JNSI
DOI: 10.24014/jnsi.v9i1.39096

Measuring Changes in Students' Science Literacy through E-Module-Based Physics Learning Using Rasch Analysis

Abdurrohman Khotim Nugraha[1*], Chaerul Rochman[2], Dindin Nasrudin[2], Fajriati Ramadhini Nur Mileniati[3], Siti Rahaimah binti Ali[4], Sitti Maesuri Patahuddin[5], Mosaddek Hossain[6], Dato' Norsabrina Mohd Noor[7]

[1] *Department of Physics, Universitas Negeri Semarang, Indonesia*
[2] *Department of Physics Education, UIN Sunan Gunung Djati, Indonesia*
[3]*Department of Nutrition Science, UIN K.H. Abdurrahman Wahid, Indonesia*
[4]*Department of Mathematics Education, Universiti Pendidikan Sultan Idris, Malaysia*
[5]*Department of Mathematics Education, University of Canberra, Australia*
[6]*Kidzana Creative Centre, Bangladesh*
[7]*World Youth Foundation, Malaysia*

**Correspondence Author: Khotim.abdurrohman@gmail.com*

ABSTRACT
This study aims to measure changes in students' scientific literacy in physics learning through e-modules focused on momentum and impulse, using the stacking and racking Rasch model techniques. The study used a one-group pretest-posttest design involving 30 10th-grade students at a high school in Bandung. Data were analyzed using the stacking and racking Rasch model techniques to identify changes in students' scientific literacy and in item difficulty levels on the logit scale. Stacking analysis results showed that e-module-based learning could increase students' scientific literacy by 1.90 logits, while racking results indicated a decrease in item difficulty levels by -1.90 logits, suggesting that the improvement in scientific literacy directly affects how students respond to the measurement instruments. These findings indicate that integrating stacking and racking techniques provides a comprehensive view of the effectiveness of e-module-based physics learning, both in terms of changes in students' abilities and the functioning of test items. This study confirms that the Rasch model is a relevant approach for examining changes in science literacy in physics learning.

Keywords: *e-module, physics, Rasch model, scientific literacy, science education*

INTRODUCTION

Scientific literacy refers to the ability to recognize, identify, and explain various scientific phenomena that occur in the world around us. This skill empowers individuals to analyze information critically and make informed decisions based on facts and solid scientific evidence (OECD, 2023). By being scientifically literate, people can navigate complex issues and contribute thoughtfully to discussions about science-related topics in their daily lives. This competency plays an important role in shaping individuals who can make accurate and rational decisions while also developing critical thinking, problem-solving, innovation, creativity, and adaptability to the development of science and technology (Cahyaningtyas, 2024; Costa et al., 2021; Sari et al., 2022). Therefore, scientific literacy becomes a fundamental competence that students must possess to adapt to the dynamics of science and technology and participate in responsible, sustainability-oriented decision-making. However, various empirical findings indicate that the level of scientific literacy among students in Indonesia is not yet optimal (Budiwati et al., 2025; Bungawati, 2024; Takda et al., 2023; Zulaiha & Meisadewi, 2022).

A recent report from the Organisation for Economic Co-operation and Development (OECD), utilizing data from the Program for International Student Assessment (PISA), reveals a concerning trend in the science literacy of Indonesian students. Over the past decade, their scores have not only declined but have consistently remained below the global average, highlighting significant challenges in the educational landscape of Indonesia (OECD, 2023). This nationwide trend aligns with preliminary findings from a public high school in Bandung. Based on the results of the students' science literacy performance tests, it was identified that their science literacy achievement is low. This situation indicates that science education, particularly in schools, has not yet fully fostered students' science literacy. Learning strategies are needed to strengthen science literacy, one of which is learning via electronic modules (e-modules) as an interactive, adaptive physics learning medium tailored to students' diverse characteristics and learning styles (Kristina et al., 2022). Several studies have shown that the use of e-modules in physics instruction improves students' science literacy (Andriani & Masykuri, 2021; Kurniawati et al., 2023; Nasrudin, Rochman, Zakiyah, et al., 2024).

Measurement of changes in students' science literacy is generally conducted using the n-gain (normalized gain) and classical statistical tests (Andriani & Masykuri, 2021; Kurniawati et al., 2023; Nasrudin, Rochman, Zakiyah, et al., 2024; Serepinaha et al., 2026; Syamsudin et al., 2026). However, both tests have limitations, namely that their scores are ordinal. The n-gain test is highly dependent on pretest scores (pretest gain correlation), which can lead to measurement bias for individuals who have high pretest achievements (McKagan et al., 2022; Nissen et al., 2018). In addition, classical statistical tests based on group-mean analysis tend to obscure individual differences in ability (Erfan et al., 2020; Jafar et al., 2024). These limitations can lead to interpretive bias and inaccurate information about changes in students' science literacy. Therefore, another measurement method is needed that can accommodate the limitations of the n-gain test and classical statistical tests, namely the Rasch model.

The Rasch model is a measurement method that analyzes test scores based on the relationship between an individual's ability and the difficulty of test items, expressed in logit units (Linacre & Wright, 1993). This unit is obtained by transforming ordinal data into interval data, allowing for the accurate measurement of changes in individual ability and avoiding measurement bias (Linacre & Wright, 1995). In measuring changes in students' science literacy, the Rasch model's stacking and racking techniques are used to identify changes in students' abilities and shifts in item difficulty levels before and after learning interventions (Linarce, 1995). Stacking reflects changes in science literacy through positive changes in logit values, while racking indicates changes in item difficulty through negative changes in logit values (Laliyo et al., 2022; Wright, 2003). The combination of both provides a comprehensive indicator of learning effectiveness.

Although the Rasch model has been used in physics education to assess the validity of test instruments (Maqruf et al., 2025; Samsudin et al., 2023) and instructional materials (Nurdini et al., 2025), a recent literature review indicates a significant research gap. Most physics education research still measures learning effectiveness using classical statistical tests or n-gain tests, which are prone to bias (Nasrudin, Rochman, & Diningsih, 2024; Nugraha et al., 2022; Solehayati et al., 2025). Additionally, stacking and racking techniques have been used in previous studies to measure changes in conceptual understanding (Laliyo et al., 2022; Utami et al., 2025) and students' learning experiences (Ramadhani et al., 2024). The use of stacking and racking techniques has not yet been specifically examined to measure changes in science literacy.

Recognizing a significant gap in existing research, this study introduces a novel approach by employing advanced Rasch model stacking and racking techniques. These methods are utilized to thoroughly investigate the longitudinal changes in students' science literacy over time. By analyzing these dynamics, the study effectively differentiates between variations in individual students' abilities and fluctuations in the difficulty levels of the assessment items. This distinction allows for a more nuanced understanding of how students evolve in their science literacy skills and how the challenges presented by the assessment itself may shift over time. Therefore, this study aims to measure changes in students' science literacy in e-module-based physics learning on the topics of momentum and impulse. It is hoped that the findings of this study will provide diagnostic solutions to address gaps in physics education evaluation methods and serve as a practical foundation for the reliability of the Rasch model as a more precise alternative to classical statistics in measuring students' science literacy and the quality of assessment instruments.

METHODOLOGY

To measure changes in students' science literacy through e-module-based physics instruction using the stacking and racking techniques of the Rasch model, this study employs a quantitative, one-group pre-posttest pre-experimental design (Creswell, 2014). This research design was specifically chosen because it allows for the longitudinal assessment of changes in students' science literacy before and after the intervention within a single learning group, without involving a comparison group, within a consistent measurement framework (Fraenkel & Wallen, 2011). The intervention was conducted over two 90-minute sessions (a total of 180 minutes), during which all students received the same physics instruction through an e-module on momentum and impulse.

This study took place at a public high school in Bandung, West Java. The participants were all 10th-grade students in the MIPA (Mathematics, Physics, and Chemistry) track, which included 12 classes. The sample was selected using purposive sampling, yielding a single class as the study sample, comprising 30 students (10 males and 20 females). This class was selected based on a specific characteristic: it had the lowest average physics grade among all other classes. Selecting a class with this characteristic was highly relevant because it provided an optimal baseline for measuring, evaluating, and contrasting the effectiveness of using e-modules in stimulating improvements in science literacy. In addition to students as research subjects, this study also involved two Physics Education lecturers from a state Islamic university in Bandung, who served as e-module validators to ensure the suitability of the learning instruments before the intervention was carried out.

The research instruments consist of an e-module designed as a medium for physics education and a science literacy assessment tool. The e-module underwent a preliminary feasibility evaluation, conducted by both media and content experts, to validate its effectiveness as an educational resource. The science literacy assessment comprises 12 multiple-choice questions, each offering five answer options, which are aligned with the indicators set forth in the PISA 2018 framework for science literacy. The item distribution for the science literacy test is detailed in Table 1.

Table 1. Distribution of Test Item Questions for Science Literacy Instrument

Competency	Competency Indicator	Question Number
Explaining scientific phenomena	Recalling and applying appropriate scientific knowledge	1
	Identifying and creating explanatory models and representations	2
	Making and justifying correct predictions	3
	Formulating clear hypotheses	4
Designing and evaluating scientific investigations	Identifying problems that can be investigated and studied scientifically	5
	Distinguishing problems that can be studied scientifically	6
	Explaining and evaluating to ensure the reliability of scientific data/information and objects	7

Competency	Competency Indicator	Question Number
	Explaining and evaluating to ensure the reliability of scientific data/information and objects	8
Interpreting scientific data and evidence	Interpreting data from one form of representation to another	9
	Analyzing and interpreting data to draw accurate conclusions	10
	Distinguishing arguments based on scientific facts and evidence from arguments based on other considerations	11
	Evaluating scientific arguments	12

The feasibility analysis of the e-module and the quality of the science literacy test instruments in this study were conducted using the Rasch model with the same validity criteria. The instruments are considered valid if they meet: 1) variance explained by Rasch measures (VERM) of at least 20%; 2) Mean Square (MnSq) value of 0.5 – 1.5; 3) Z-Standard (ZStd) value from -2.0 to +2; 4) Point Measure Correlation (PtMea Corr) value of 0.4 – 0.85 (Boone et al., 2014; Sumintono, 2015). The results of the feasibility test of the physics learning e-module on the topic of momentum and impulse are presented in Table 2

Table 2. E-Module Feasibility Test Results

Feasibility Aspect	Component Test	Assessment Indicator	Result	Criteria	Description
Display	Dimensionality	Variance explained by Rasch measures (VERM)	53%	Unidimensionality	Worthy
	Validity	Outfit MnSq	0,6	Valid	
		Outfit ZStd	0,2	Valid	
		Outfit PtMea Corr	0,7	Valid	
Content	Dimensionality	Variance explained by Rasch measures (VERM)	85%	Unidimensionality	Worthy
	Validity	Outfit MnSq	0,5	Valid	
		Outfit ZStd	0,5	Valid	
		Outfit PtMea Corr	0,8	Valid	

Meanwhile, the results of the instrument quality assessment for the science literacy test are presented in Table 3 below.

Table 3. Results of the Science Literacy Instrument Quality Test

Component Test	Indicator	Results	Criteria	Description
Dimensionality	Variance explained by Rasch measures (VERM)	22,5%	Unidimensionality	Can measure one latent construct, namely scientific literacy
Validity	Outfit MnSq	0,7 – 1,6	Valid	The test instrument is able to measure scientific literacy
	Outfit ZStd	-1,7 – 2,4	Valid	
	Outfit PtMea Corr	0,2 – 0,6	Valid	
Reliability	Item reliability	0,8	Good	The test instrument has good consistency in measuring scientific literacy
	Cronbach Alpha (KR-20)	0,9	Good	

Research data were collected through pre- and posttests to measure students' science literacy before and after the learning intervention. Students completed the test instruments on the provided answer sheets under the teacher's supervision for 45 minutes. All student response data were collected directly and compiled after the test. The research data analysis was conducted using the Rasch model with MINISTEP 5.10.4.0. The pretest and posttest data, initially ordinal, were converted into interval data in logit units. Changes in students' science literacy were measured using the stacking technique, while changes in item characteristics were analyzed using the racking technique. An increase in science literacy is indicated by a positive shift in students' ability logit scores on the posttest. In contrast, changes in item characteristics are shown by a negative shift in item difficulty logit scores. Both techniques allow the analysis of changes in students' abilities and item characteristics to be conducted within a consistent measurement framework.

The research procedure began with the preparation of a physics e-module on momentum and impulse, as well as a science literacy test instrument. The e-module was first tested for feasibility, and the science literacy test instrument was evaluated for quality. After the e-module was deemed feasible and the instrument met the measurement quality criteria, the research proceeded with a pretest to assess the students' initial science literacy. Next, the physics learning intervention based on the e-module was carried out. After the intervention, the students were given a posttest using the same instrument as the pretest. The pretest and posttest data were then analyzed using stacking and racking techniques to measure changes in students' science literacy. The research procedure flowchart is presented in Figure 1.

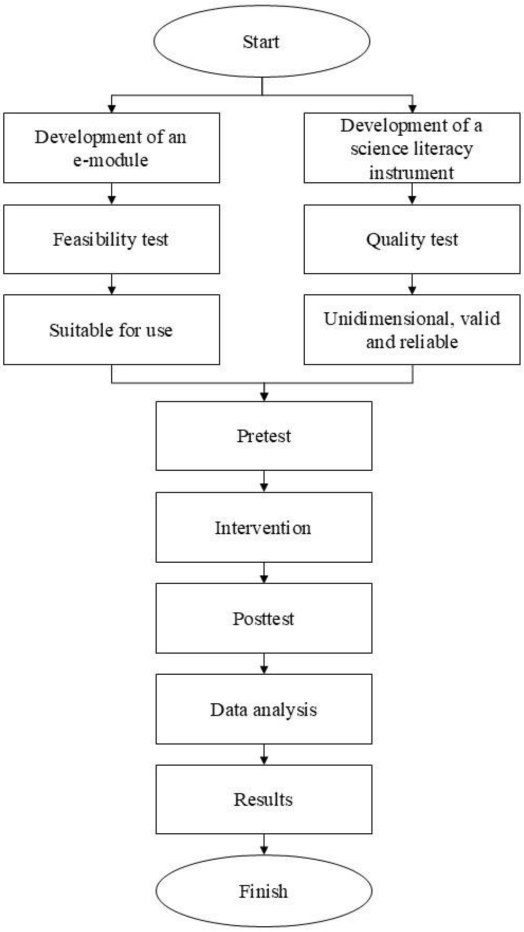

Figure 1. Research Flow Diagram

RESULT AND DISCUSSION

In line with the objective of this study, which is to explicitly measure changes in students' science literacy in e-module-based physics learning on the topics of momentum and impulse using the stacking and racking techniques of the Rasch model, the results of the analysis provide comprehensive empirical evidence. The logit calculation results from these two techniques directly address the research objectives by dissecting students' science literacy levels under two measurement conditions namely, before and after the learning intervention as detailed below.

The stacking technique was used to measure changes in students' science literacy across two measurement conditions: pretest (before the intervention) and posttest (after the intervention). Table 4 presents the average pretest and posttest scores for students' science literacy, along with the logit differences from the Rasch model analysis.

Table 4. Average Science Literacy Score of Students

Number of Students	Average Pretest Score (logit)	Average Posttest Score (logit)	Difference of Average Posttest-Pretest (logit)
30	0,33	2,23	1,90

Based on Table 4, the average science literacy score of 30 students before the e-module-based physics learning intervention was 0.33 logits, while the average score after the intervention was 2.33 logits. The difference in the average of 1.90 logits indicates an improvement in students' science literacy after the learning intervention. Meanwhile, changes in each student's science literacy are shown in the pre-posttest scatter plot in Figure 2.

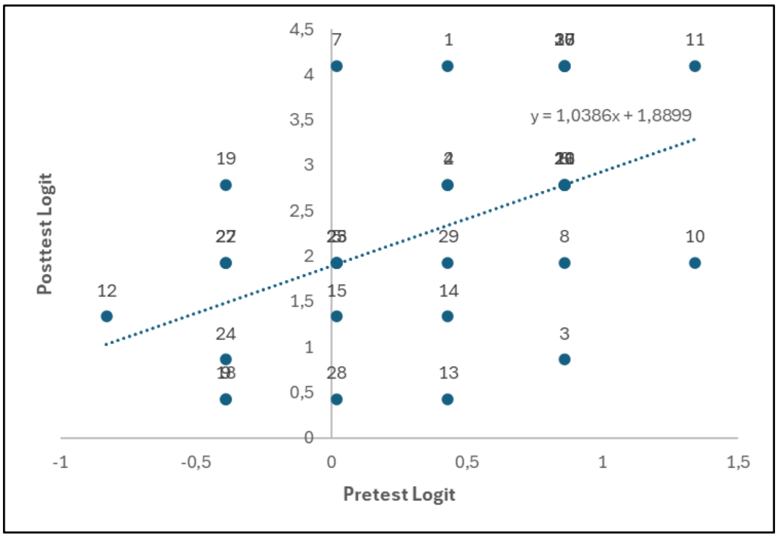

Figure 2. Scatter Plot Graph of Students' Science Literacy Pre- and Post-Test

Based on Figure 2, students are represented by blue dots on the scatter plot; the equation of the line defines the boundary for changes in students' science literacy. On the vertical axis, the size of science literacy results from the posttest is in the range of 0 to 4.5 logits. In contrast, on the horizontal axis, the size of science literacy results from the pretest is in the range of -1 to 1.5 logits. The analysis showed that 19 students exceeded the change threshold, with student 7 having the largest pre-posttest logit difference (>4 logits), indicating the greatest increase in science literacy after the e-module-based physics learning intervention. Meanwhile, 11 students were below the change boundary line, with students 3 and 13 showing a minimal pre-posttest logit difference, indicating that the increase in science literacy was relatively insignificant.

The changes observed in science literacy, which were identified using stacking techniques, can be understood as a direct result of physics learning interventions. These interventions utilized electronic modules specifically designed to address the concepts of momentum and impulse, highlighting their effectiveness in enhancing students' understanding of these fundamental topics in physics. The e-module presents the concepts of momentum and impulse through learning activities that require students to identify phenomena in everyday life contexts, conduct simple experiments, analyze observation and measurement data, and draw conclusions based on the obtained evidence. These learning activities align with the core competencies of science literacy. This study's results are consistent with research by Susilawati et al., (2025) and Kates et al., (2026) which reported that the use of activity- and context-based e-modules can improve students' science literacy. Thus, the stacking results confirm that physics learning through e-modules leads to functional and contextual changes in students' science literacy, rather than merely improving learning achievement.

Although e-module-based learning generally shows a positive effect on science literacy, students' positions relative to the change threshold indicate heterogeneous responses. As many as 19 students are above the change line, indicating their ability to engage in activities such as phenomenon identification, experimentation, and data analysis to improve science literacy. Meanwhile, 11 students are below the threshold, which does not entirely reflect the ineffectiveness of learning but is influenced by several factors, including limitations in scientific reasoning (Gardner et al., 2024), difficulty interpreting data (Ghaleb & Vui, 2025), low learning independence and limited time (Ghaleb & Vui, 2025), also teacher competence (Ekmekci & Serrano, 2022). On the other hand, students with high initial science literacy, such as students 3 and 13, showed no significant improvement due to the ceiling effect (Garin, 2014; Medvedev & Krägeloh, 2025).

The racking technique was applied to assess the difficulty of test items under two conditions: before and after the learning intervention. Through this technique, changes in the difficulty level of test items can be identified as students' science literacy evolves over the course of the intervention. A summary of the average test item difficulty in the pretest and posttest, along with the differences in logit scores from the Rasch model ranking analysis, is presented in Table 5 below.

Table 5. Average Item Difficulty Level

Number of Questions	Average Score Difficulty Level of Questions during Pretest (logit)	Average Score Difficulty Level of Questions during Posttest (logit)	Average Difference Posttest Pretest (logit)
12	0,63	-1,27	-1,90

Table 5 illustrates the changes in the difficulty level of science literacy test items after the implementation of e-module-based physics learning. The average difficulty level of 12 test items in the initial measurement (pretest) was 0.63 logits, then shifted to -1.27 logits in the final measurement (posttest). This average shift of -1.90 logits indicates a decrease in test item difficulty, suggesting an improvement in students' science literacy after the learning intervention. Meanwhile, the changes in the difficulty level of each test item are shown in the scatter plot graph in Figure 3 below

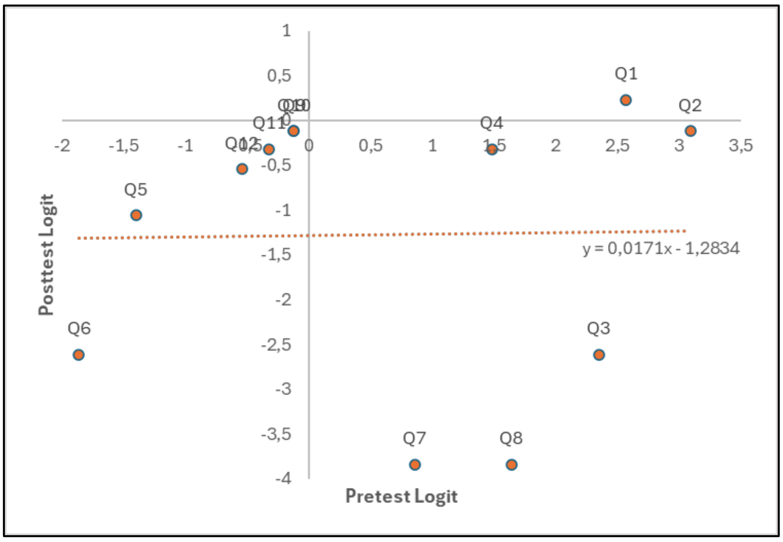

Figure 3. Scatter Plot Graph of Pre-Posttest Item Difficulty Levels

The scatter plot illustrated in Figure 3 vividly displays the variations in the difficulty levels of science literacy items, highlighting how these levels change and relate to one another.. An orange dot represents each item. The equation line serves as a reference for identifying shifts in item difficulty between the pretest and posttest. The range of difficulty levels on the posttest, shown on the vertical axis, is -4 to 1 logits, while those on the pretest, shown on the horizontal axis, range from -2 to 3.5 logits. The analysis shows that 4 items, namely Q3, Q6, Q7, and Q8, are below the reference line, indicating a decrease in difficulty levels after the physics learning intervention based on e-modules. Conversely, the other 8 items are above the reference line, indicating that the decrease in difficulty was not very significant.

The application of the racking technique enables systematic interpretation of changes in item difficulty between pretest and posttest conditions within a consistent measurement scale. The shift in item difficulty after instructional intervention indicates that improvements in students' science literacy directly affect how students respond to the measurement instrument (Laliyo et al., 2022). In the context of physics learning through e-modules on momentum and impulse, activities such as phenomenon identification, simple experiments, and data analysis not only enhance students' science literacy but also make tasks that previously required high-level reasoning easier to solve. This aligns with research by Gardner et al., (2024); Pertiwi et al., (2024); Suwarna & Zulfiani, (2024). Thus, racking indicates that changes in the difficulty level of test items are not due to changes in the instrument's characteristics, but instead to increased students' science literacy. These findings affirm that racking complements stacking by providing a two-way perspective, namely, changes in students and in test items within a unified measurement framework.

Although the test items generally became easier after the learning intervention, the racking analysis showed that 8 of 12 items decreased in difficulty. However, the change was not significant because they were above the change line. This condition is observed in items Q1, Q2, Q3, and Q4, which already had high pretest response rates. These three items measure competence in explaining scientific phenomena, which are generally easier for students to understand because they relate to everyday life contexts, thereby limiting the increase in difficulty. This is known as the ceiling effect (Garin, 2014; Medvedev & Krägeloh, 2025). In contrast, item Q5, which measures competence in designing and evaluating scientific investigations, shows the smallest decline in difficulty. This is due to the cognitive complexity of this competence, which requires the ability to design experimental procedures, identify variables, and evaluate measurement data, thus necessitating guidance and supervision from teachers during experiments (Szalay et al., 2023; van Riesen et al., 2018).

The results of the discussion highlight a significant enhancement in students' science literacy through the incorporation of e-modules in physics education, specifically focusing on the concepts of momentum and impulse. This improvement is evident not only in students' overall capabilities but also in the characteristics of the measurement instruments used to assess their learning. A stacking analysis further reveals a notable transition toward higher levels of science literacy among the students. Moreover, a racking analysis indicates that as science literacy improves, there is a corresponding decrease in the complexity of assessment items. This suggests that the test questions become more accessible, allowing students to respond with greater ease. Overall, the findings underscore the positive impact of e-modules on educational outcomes in the realm of physics. These findings are in line with Amin et al., (2017) and Khine, (2020), who affirmed the Rasch model's ability to objectively separate changes in respondents' abilities from the instrument's characteristics. The integration of stacking and racking provides a two-way perspective on the effectiveness of learning interventions, namely the development of students' science literacy and changes in item difficulty within a complete measurement framework (Bode & Wright, 1999). Thus, this study reinforces stacking and racking as valid and comprehensive analytical approaches to examine improvements in science literacy in physics learning.

The significance of this study lies in its innovative approach to providing a precise diagnostic tool for evaluating the effectiveness of science education, particularly in physics. It aims to address the limitations and biases inherent in traditional statistical testing methods. The study's novelty is showcased through the application of advanced Rasch model stacking and racking techniques, specifically used to analyze the intricate dynamics of science literacy within a robust, comprehensive measurement framework. By leveraging these methodologies, the research establishes a new, objective framework that educators and researchers can use to discern genuine changes in students' capabilities, distinguishing them from mere fluctuations in item difficulty. This endeavor has significant practical implications, as it encourages policymakers to adopt Rasch model-based evaluations for assessing science literacy. Additionally, it recommends the use of e-modules that incorporate contextual experimental activities, which are validated and effective learning tools.

Nevertheless, it is essential to acknowledge the study's limitations. These include the relatively brief intervention duration, the lack of a control group to allow a direct comparison of outcomes, and the focus on a single school with a low baseline in science literacy among its students. To strengthen future research, it is strongly advised to implement a quasi-experimental design with a control group, extend the intervention duration, and increase the sample size to include participants from diverse backgrounds and demographics. This would enhance the reliability of the findings and help confirm their consistency across varied contexts.

CONCLUSION

In response to the research objectives, it can be concluded that implementing e-module-based physics instruction on momentum and impulse has been proven to effectively and measurably improve students' science literacy. This enhancement goes beyond a simple increase in scores; it can be thoroughly understood through a detailed analysis of the stacking and racking techniques utilized within the Rasch model. These methods not only contribute to more accurate measurements but also provide a nuanced insight into the underlying patterns and relationships among the data. The stacking analysis results show a positive shift in student ability of +1.90 logits, indicating meaningful cognitive reconstruction, with contextual experimental activities within the e-module successfully facilitating better problem solving of scientific phenomena. This scientific significance is closely aligned with the racking results, which show a negative shift in item difficulty of -1.90 logits. This decrease in difficulty level demonstrates that improved science literacy directly breaks down cognitive barriers in test items, making previously high-reasoning-demand instruments easier to solve thanks to the development of students' cognitive schemas. In this context, the use of stacking and racking techniques demonstrates their sharp methodological contribution not merely as a measurement metric, but as a comprehensive diagnostic tool that presents a two-way picture of students' abilities and item functions integrated within a single, consistent measurement system.

REFERENCES

Amin, N. F., Latif, A. A., Arsat, M., & Suhairom, N. (2017). Measuring changing of problem solving ability using Rasch measurement model. *Advanced Science Letters*, *23*(1), 117–120. https://doi.org/10.1166/asl.2017.7264

Andriani, S., & Masykuri, M. (2021). Development of electronic modules (e-modules) based on guided inquiry on temperature and heat materials to improve students' science literacy. *Jurnal Penelitian Pendidikan IPA*, *7*(SpecialIssue), 281–287. https://jppipa.unram.ac.id/index.php/jppipa/article/view/1234

Bode, R. K., & Wright, B. D. (1999). Rasch measurement in higher education. In *Higher education: Handbook of theory and research* (pp. 287–316). Springer.

Boone, W. J., Yale, M. S., & Staver, J. R. (2014). Rasch analysis in the human sciences. In *Rasch Analysis in the Human Sciences*. https://doi.org/10.1007/978-94-007-6857-4

Budiwati, S., Nugroho, A. S., & Hayat, M. S. (2025). *Profil kemampuan literasi sains pada peserta didik fase E MAN Kota Sorong Papua Barat Daya tahun 2025*. *13*(1), 68–76.

Bungawati, B. (2024). Profil kemampuan literasi sains siswa pada materi ekosistem. *Jurnal Binomial*, *7*(1), 43–50. https://doi.org/10.46918/bn.v7i1.2185

Cahyaningtyas, R. (2024). *Student's science literacy in science learning at elementary school*. *15*, 1–14.

Costa, A., Loureiro, M., & Ferreira, M. E. (2021). Scientific literacy: the conceptual framework prevailing over the first decade of the twenty-fisrt century. In *Revista Colombiana de Educación* (Vol. 1, Issue 81). https://revistas.upn.edu.co/index.php/RCE/article/view/10293

Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (4th Edition). SAGE Publications, Inc.

Ekmekci, A., & Serrano, D. M. (2022). The Impact of Teacher Quality on Student Motivation, Achievement, and Persistence in Science and Mathematics. *Education Sciences*, *12*(10). https://doi.org/10.3390/educsci12100649

Erfan, M., Maulyda, M. A., Hidayati, V. R., Astria, F. P., & Ratu, T. (2020). *Analisis kualitas soal kemampuan membedakan rangkaian seri dan paralel melalui teori tes klasik dan model Rasch*. *3*(1), 11–19.

Fraenkel, J., & Wallen, N. (2011). *How to Design and Evaluate Research in Education (8th Ed.)*. McGraw-Hill Publishing. https://books.google.co.id/books?id=bUOPtQEACAAJ

Gardner, S. M., Angra, A., & Harsh, J. A. (2024). Supporting Student Competencies in Graph Reading, Interpretation, Construction, and Evaluation. *CBE Life Sciences Education*, *23*(1). https://doi.org/10.1187/cbe.22-10-0207

Garin, O. (2014). *Ceiling Effect BT - Encyclopedia of Quality of Life and Well-Being Research* (A. C. Michalos (ed.); pp. 631–633). Springer Netherlands. https://doi.org/10.1007/978-94-007-0753-5_296

Ghaleb, M. M. S., & Vui, C. N. (2025). Mindset Assessment Profile and Behavior Change: Role of Self-Regulated Learning Behavior and Thought Content and Control. *Journal of Mind and*

Behavior, *46*(2), 434–461. https://www.scopus.com/inward/record.uri?eid=2-s2.0-105025793274&partnerID=40&md5=63bf1e787a79ae67e5ec74971119edb2

Jafar, L., Riyadi, & Ridwan, A. (2024). *Penggunaan teori tes klasik untuk analisis butir soal pada asesmen*. *02*(03), 1–10.

Kates, M., Fitzpatrick, A., Zhang, Y., Gibbs, A., Liu, Q., da Silveira Fleck, A., Freedman, L., & Andreopoulos, S. (2026). Effectiveness of Scientific Literacy and Communication E-Modules in a Molecular Biology Course: A Qualitative and Quantitative Analysis. *Biochemistry and Molecular Biology Education*. https://doi.org/10.1002/bmb.70043

Khine, M. S. (2020). Objective measurement in psychometric analysis. In *Rasch Measurement: Applications in Quantitative Educational Research* (pp. 3–7). https://doi.org/10.1007/978-981-15-1800-3_1

Kristina, H., Vitasari, M., & Taufik, A. N. (2022). Pengembangan e-modul berbasis literasi sains tema ayo siaga bencana untuk melatih kemandirian belajar siswa SMP. *PENDIPA Journal of Science Education*, *6*(3), 754–763. https://doi.org/10.33369/pendipa.6.3.754-763

Kurniawati, T. D., Akhdinirwanti, R. W., & Fatmaryanti, S. D. (2023). Pengembangan e-modul menggunakan aplikasi 3D PageFlip Professional untuk meningkatkan kemampuan literasi sains peserta didik. *Jurnal Inovasi Pendidikan Sains (JIPS*, *4*(1), 10–16.

Laliyo, L. A. R., Sumintono, B., & Panigoro, C. (2022). Measuring changes in hydrolysis concept of students taught by inquiry model: stacking and racking analysis techniques in Rasch model. *Heliyon*, *8*(3), e09126. https://doi.org/10.1016/j.heliyon.2022.e09126

Linacre, J. M., & Wright, B. (1995). *Rasch Measurement Transactions*. MESA Press Chicago, IL, USA:

Linacre, J. M., & Wright, B. D. (1993). Constructing linear measures from counts of qualitative observations. *International Conference on Bibliometrics, Informetrics and Scientometrics*. http://www.winsteps.com/a/Linacre-Wright-Constructing.pdf

Linarce, J. M. (1995). *Rasch Measurement Transactions*. MESA Press Chicago.

Maqruf, A., Suhandi, A., Samsudin, A., Sözbilir, M., & Özgür, H. (2025). *Exploring students' creative thinking skills: a design of reliable instrument in term of assessing creative thinking on temperature and heat concepts*. *8*(2), 258–272. https://doi.org/10.24014/jnsi.v8i2.31857

McKagan, S., Sayre, E., & Madsen, A. (2022). *Normalized gain: What is it and when and how should I use it?* https://www.physport.org/recommendations/Entry.cfm?ID=93334

Medvedev, O. N., & Krägeloh, C. U. (2025). Rasch measurement model. In *Handbook of assessment in mindfulness research* (pp. 131–147). Springer.

Nasrudin, D., Rochman, C., & Diningsih, I. J. (2024). *Developing biogas digital comic for improving renewable energy literacy*. *2024*, 407–414. https://doi.org/10.18502/kss.v9i8.15572

Nasrudin, D., Rochman, C., Zakiyah, R., & Irwansyah, F. S. (2024). Development of learning media based on augmented reality technology to improve student's scientific literacy in the context of ocean energy. *AIP Conference Proceedings*, *3058*(1), 20008.

Nissen, J. M., Talbot, R. M., Thompson, A. N., & Dusen, B. Van. (2018). Comparison of normalized gain and Cohen ' s d for analyzing gains on concept inventories. *Physical Review Physics Education Research*, *14*(1), 10115. https://doi.org/10.1103/PhysRevPhysEducRes.14.010115

Nugraha, A. K., Rochman, C., & Nasrudin, D. (2022). Senior High School Students' Literacy Profile on Energy Conversion Process. *Jurnal Riset Pendidikan Fisika*, *7*(1), 29.

Nurdini, N., Suhandi, A., Fratiwi, N. J., & Samsudin, A. (2025). *E-CDCCText development: an instructional material for facilitating conceptual change in fluid concepts*. *8*(1), 93–110. https://doi.org/10.24014/jnsi.v8i1.32928

OECD. (2023). *PISA 2025 SCIENCE FRAMEWORK*. Oxford University Press.

Pertiwi, B., Purwanto, A., & Setiawan, I. (2024). Development of A Problem Based Learning Oriented Physics E-Module to Improve The Scientific Literacy of High School Students. *Jurnal Pendidikan Matematika Dan IPA*. https://api.semanticscholar.org/CorpusID:267058395

Ramadhani, D. G., Yamtinah, S., Saputro, S., Sarwanto, S., & Masykuri, M. (2024). From screen to bench: unpacking the shifts in chemistry learning experiences during the covid-19 transition. *Chemistry Teacher International*, *6*(1), 19–33. https://doi.org/10.1515/cti-2023-0039

Samsudin, A., Aminudin, A. H., Novia, H., & Suhandi, A. (2023). *Identifying javanese students' conceptions on fluid pressure with wright map analysis of rasch*. *6*(2), 173–185. https://doi.org/10.24014/jnsi.v6i2.21822

Sari, S. W., Yokhebed, & Tenriawaru, A. B. (2022). Analisis kemampuan literasi sains pada materi virus kelas X di SMA Negeri 1 Rasau Jaya. *Jurnal Pendidikan MIPA*, *12*(1), 1–7.

Serepinaha, M., Rahmawati, Y., & Wibowo, F. C. (2026). Integrating STEAM-EDP and design thinking to enhance scientific literacy: A design-based research in Indonesian elementary schools. *Multidisciplinary Science Journal*, *8*(8). https://doi.org/10.31893/multiscience.2026525

Solehayati, R., Kariadinata, R., Chusni, M. M., Islam, U., Sunan, N., & Djati, G. (2025). *Penerapan model pembelajaran inquiry berbantuan simulasi PhET untuk meningkatkan literasi sains peserta didik kelas XII*. *3*(3), 123–133.

Sumintono, B. (2015). *Aplikasi Pemodelan Rasch pada Assessment Pendidikan Knowledge Management View project seven system leadership study View project*. *155*. www.trimkomunikata.com

Susilawati, S., Mulyani, G., Halim, A., Saputri, M., Farhan, A., & Widia, S. (2025). Enhancing Scientific Literacy through an Interactive E-Module Enriched with Acehnese Local Wisdom. *Journal of Physics: Conference Series*, *3148*(1). https://doi.org/10.1088/1742-6596/3148/1/012008

Suwarna, I. P., & Zulfiani, Z. (2024). Development of WIV-S Physics e-Learning to Improve Inquiry Abilities and Digital Literacy of Prospective Science Teacher Students. *International Journal of Information and Education Technology*, *14*(9), 1291–1298. https://doi.org/10.18178/ijiet.2024.14.9.2159

Syamsudin, F. I., Saputro, S., Sarwanto, S., Karyanto, P., & Kamari, A. (2026). Effectiveness of Problem and Design Process-Based Learning Model with STEAM Approach to Improve Students' Scientific Literacy in Science Learning in Rural and Urban Areas. *Ianna Journal of Interdisciplinary Studies*, *8*(1), 242–256. https://doi.org/10.5281/zenodo.17700295

Szalay, L., Tóth, Z., Borbás, R., & Füzesi, I. (2023). Scaffolding of experimental design skills. *Chemistry Education Research and Practice*, *24*(2), 599–623. https://doi.org/10.1039/d2rp00260d

Takda, A., Arifin, K., & Tahang, L. (2023). Profil kemampuan literasi sains peserta didik SMA berdasarkan Nature of Science Literacy Test (NoSLiT). *Jurnal Penelitian Pendidikan Fisika*, *8*(1), 19–27. https://doi.org/10.36709/jipfi.v8i1.7

Utami, N., Setiawan, A., Hamidah, I., & Koehler, T. (2025). Investigating problem-based worksheets (PBWs) to improve understanding in logic gates topic: stacking and racking analyses in rasch model. *Educational Process: International Journal*, *16*. https://doi.org/10.22521/edupij.2025.16.243

van Riesen, S., Gijlers, H., Anjewierden, A., & de Jong, T. (2018). Supporting learners' experiment design. *Educational Technology Research and Development*, *66*(2), 475–491. https://doi.org/10.1007/s11423-017-9568-4

Wright, B. D. (2003). *Rack and Stack: Time 1 vs Time 2 or Pre-Test vs Post-Test.* https://www.rasch.org/rmt/rmt171a.htm

Zulaiha, F., & Meisadewi, N. (2022). Profil literasi sains siswa SMA kelas X pada masa pandemi COVID-19. *Aswaja*, *3*(1), 66–75.