Journal of Psychology and Culture Behavior in SDGs
Homepage : https://journal.i-ros.org/index.php/JOPACBIS
Email : jopacbis@i-ros.org

e-ISSN 3110-1615
Vol. 2, No.2, July 2026

Development of an Integrated Science Literacy Assessment Instrument Incorporating Local Gamelan Wisdom in the Topic of Sound Waves to Support the Achievement of SDG 4: Quality Education

Ayunda Zahra Aulia[1*], Titin Sunarti[1], Joshi Maharani Wibowo[2]
[1*]Universitas Negeri Surabaya, Surabaya, Indonesia
[2*]University of Helsinki, Finland

Check for updates OPEN ACCESS

DOI : https://doi.org/10.63230/jopacbis.2.2.249

Sections Info

Article history:
Submitted: June 12, 2026
Final Revised: June 30, 2025
Accepted: June 30, 2025
Published: July 01, 2026

Keywords:
Science literacy
Local gamelan wisdom
Sound waves
Assessment instrument

ABSTRACT

Objective: *This study aims to develop an essay-based science literacy assessment instrument that integrates local gamelan wisdom into the topic of sound waves for high school students, in accordance with the PISA 2025 science literacy framework.* ***Method:*** *The ADDIE (Analysis, Design, Development, Implementation, and Evaluation) development model was used, limited to the development stage. The instrument consists of ten essay questions, a question outline, and a scoring rubric. Validation was conducted by three experts, while the empirical test involved 35 students and was analyzed using the Rasch Model with Winsteps software.* ***Results:*** *Expert validation showed a validity score of 86% (highly valid). Rasch analysis indicated that all test items met the criteria for empirical validity, with Point-Measure Correlation values ranging from 0.36 to 0.77 and MNSQ values ranging from 0.52 to 1.35. The item reliability coefficient of 0.91 falls into the "very good" category, while the person reliability coefficient of 0.76 falls into the "adequate" category. The Wright Map shows item difficulty levels ranging from −1.20 to +4.00 logits, supported by an item separation index of 2.91, indicating the instrument's strong ability to distinguish difficulty levels.* ***Novelty:*** *The novelty of this study lies in the integration of the acoustic phenomena of the local wisdom of gamelan into a science literacy instrument in the form of an essay, which is aligned with the PISA 2025 framework and tested using the Rasch Model. This instrument has proven to be suitable for measuring students' science literacy in a contextual manner while supporting cultural preservation and the achievement of SDG 4 of Quality Education.*

INTRODUCTION

Quality Education

Quality education is a key pillar of the Sustainable Development Goals (SDGs), particularly Goal 4 (SDG 4), which emphasizes the importance of providing inclusive, equitable, and quality education to promote lifelong learning opportunities for all. In the context of the 21st century and the Merdeka Curriculum, the primary focus of learning has shifted from passive mastery of subject matter toward strengthening foundational competencies that enable students to think critically, creatively, communicatively, and collaboratively in addressing global challenges (Fauzan et al., 2023). This improvement in quality demands a fundamental overhaul of both the learning system and assessment tools in schools so that they can reflect future needs.

Science Literacy

One of the essential competencies that must be developed is science literacy the ability of individuals to use scientific knowledge to explain phenomena, evaluate information, and make decisions based on scientific evidence. This ability serves as a key indicator in the Programme for International Student Assessment (PISA) conducted by the OECD.

e-ISSN 3110-1615
Vol. 2, No.2, July 2026

Within the PISA 2025 framework, science literacy encompasses three core competencies: explaining phenomena scientifically; designing and evaluating scientific experiments; and researching, evaluating, and using scientific information for decision-making and practical action (OECD, 2023). Although important, students' science literacy achievements in Indonesia remain relatively low and consistently fall below the OECD average, according to the PISA 2022 results (OECD, 2023). This situation is exacerbated by the school assessment system, which generally remains focused on rote memorization of scientific concepts rather than the development of higher-order thinking skills (Adao & Morong, 2025). In addition, assessment practices tend to emphasize routine problem-solving, limiting opportunities for students to demonstrate scientific reasoning and analytical thinking (Heim et al., 2022). As a result, existing assessment instruments often fail to comprehensively measure students' abilities to interpret data and evaluate scientific evidence (Santoso et al., 2023). Consequently, students have limited opportunities to develop the science literacy competencies emphasized in the PISA framework (Alfarizy et al., 2025). Therefore, authentic assessment instruments that are closely aligned with students' real-world environments are needed to facilitate meaningful learning experiences. Such instruments should encourage students to connect scientific concepts with real-life situations through contextual learning experiences (Gilbert, 2006). In this way, the process of constructing scientific knowledge becomes more meaningful because learning is built upon students' interactions with their social and cultural environments (Vygotsky, 1978).

Local Wisdom of Gamelan

Integrating local wisdom into assessment instruments can serve as a strategic approach to implementing contextual assessment practices that are relevant to students' daily lives (Shufa, 2018). In addition, the incorporation of local cultural values into educational assessment contributes to preserving cultural heritage while strengthening students' national identity (Sakti et al., 2024). One cultural heritage with strong scientific relevance is the Javanese gamelan (Sukerta, 2015). From a physics perspective, various gamelan instruments, such as the saron, bonang, kenong, peking, and gong, exhibit acoustic phenomena that are closely related to the concepts of sound waves, frequency, amplitude, and resonance (Gee, 2009). These acoustic characteristics make gamelan an appropriate contextual medium for learning and assessing sound-wave concepts in physics (Yunior Erlangga et al., 2022). Recent studies have also demonstrated that the physical characteristics of gamelan instruments, including their material composition, resonance, and sound production mechanisms, provide authentic examples for explaining sound-wave concepts in physics learning (Damarsha et al., 2023).

Through the Culturally Responsive Pedagogy approach, the use of gamelan as an assessment context not only measures students' understanding of physics concepts but also encourages them to appreciate the relationship between science and their cultural heritage (Saputra et al., 2025). Such culturally relevant learning experiences promote

e-ISSN 3110-1615
Vol. 2, No.2, July 2026

meaningful knowledge construction because students learn scientific concepts within authentic social and cultural contexts (Mathis et al., 2023).

Rasch Analysis

In addition to contextual appropriateness, the theoretical and empirical quality of the assessment instrument must also be comprehensively tested. Item quality analysis in this study utilized the Rasch Model because it provides objective estimates of construct validity, reliability, and item difficulty that are independent of sample characteristics (Sumintono & Widhiastrini, 2015). Several previous studies have integrated the local wisdom of gamelan into physics learning materials to provide contextual learning experiences (Bagas Damarsha et al., 2023). Other studies have also developed learning media and instructional resources based on gamelan to enhance students' conceptual understanding and engagement in physics learning Trisiani et al. (2023). In addition, gamelan has been utilized as a contextual learning resource to facilitate students' understanding of sound-wave concepts through local cultural phenomena (Maulida & Sunarti, 2022). Recent research has continued to explore the integration of gamelan into physics education by developing culturally relevant learning innovations (Camacho-Sánchez et al., 2023). Furthermore, gamelan-based learning has been shown to improve students' learning experiences by connecting scientific concepts with local cultural contexts (Salsabilla Izzah et al., 2023). Nevertheless, most previous studies have focused primarily on learning media and instructional implementation rather than developing an essay-based science literacy assessment instrument on sound waves that is aligned with the PISA 2025 framework and validated using the Rasch Model (Abdul Raof & Musta'amal, 2021).

Research Objectives

Given this gap, this study aims to develop a science literacy assessment instrument in the form of essay questions on the topic of sound waves, integrated with the local wisdom of gamelan, for high school students (Indasa et al., 2024). The novelty of this study lies in the development of acoustic phenomenon stimuli from gamelan that are aligned with the three core competencies of the PISA 2025 framework and evaluated using the measurement characteristics of the Rasch Model (Wahyuni & Ragil Widianto Atmojo, 2024). It is hoped that this instrument will serve as a valid and reliable alternative for diagnostic assessment, support cultural preservation, and make a tangible contribution to the achievement of SDG 4 (Quality Education) in Indonesia (Adipat & Chotikapanich, 2022).

RESEARCH METHOD

This study is a development research project that employs the ADDIE model (Analysis, Design, Development, Implementation, and Evaluation). However, the study is limited to the development stage, which includes needs analysis, instrument design, expert

validation, revision, and pilot testing. The implementation and evaluation stages are planned for a follow-up study through a large-scale pilot test.

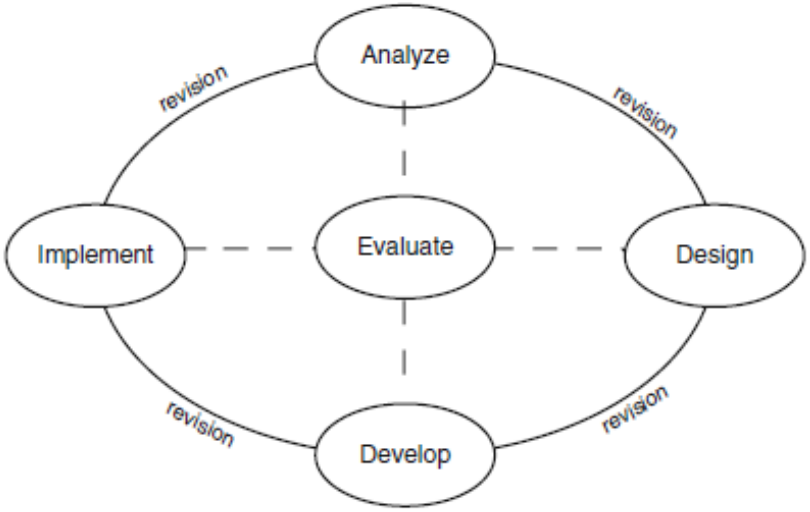

Figure 1. Flowchart of the ADDIE Research Model (Branch, 2009)

During the analysis phase, a literature review was conducted on science literacy, the PISA 2025 framework, sound wave material, and the integration of local gamelan wisdom as the assessment context. The design phase included drafting instrument specifications, developing stimuli based on gamelan acoustic phenomena, drafting 10 essay questions, and developing a scoring rubric. A summary of the instrument specifications is presented in Table 1. Instrument validation was then conducted by three validators selected through purposive sampling, consisting of two physics education lecturers and one high school physics teacher who possess expertise in the fields of learning assessment, science literacy, and assessment instrument development. Content validity was analyzed using the validity percentage and the Percentage of Agreement (PoA) to determine the level of agreement among validators, calculated using the following formula:

$$Percentage\ of\ Agreement = 100\%\ x\ \left(1 - \frac{A-B}{A+B}\right) \qquad \ldots(1)$$

Notes:
A = Frequency of aspects rated by the validator with high frequency.
B = Frequency of aspects rated by the validator as low.

Percentage of Agreement s are considered valid if they yield a PoA≥ of 75% (Borich, 1994). The pilot study involved 35 11th-grade high school students selected through purposive sampling. This sample size met the minimum recommendations for Rasch analysis during the initial pilot phase of the instrument to estimate item parameters and detect potential instrument issues before conducting a large-scale pilot test. The pilot test data were analyzed using the Rasch model with the assistance of Winsteps software to determine the instrument's empirical validity, reliability, difficulty level, and discriminative power. Linacre (1994) also stated that a sample size of approximately 30–

50 respondents is sufficient to produce relatively stable item calibration estimates during the early stages of instrument development. Furthermore, Sumintono and Widhiarso (2015) explain that the Rasch Model can be applied to small samples because the resulting parameter estimates are relatively independent of sample characteristics and still provide diagnostic information regarding the instrument's quality. Therefore, a sample of 35 students was deemed sufficient for the initial pilot testing of the instrument and the analysis of item characteristics prior to testing on a larger sample.

RESULTS AND DISCUSSION

Results

The results of the analysis show that science literacy is a critical competency in physics education and needs to be assessed using contextual instruments. The topic of sound waves is closely related to acoustic phenomena in gamelan music, making it a potential authentic context for the development of science literacy assessment instruments. Based on the analysis results, an essay-based science literacy assessment instrument was developed that aligns with the PISA 2025 science literacy competencies and integrates local gamelan wisdom. The instrument consists of 10 essay questions accompanied by a scoring rubric. A summary of the specifications of the developed instrument is presented in Table 1.

Table 1. Science Literacy Instrument Rubric

PISA 2025 Science Literacy Competencies	General Indicators	Number of Questions	Question Number
Explaining Phenomena Scientifically	Using the concept of sound waves to explain phenomena in the gamelan	4	1–4
Designing & Evaluating Scientific Inquiries and Critically Interpreting Data and Evidence	Designing and evaluating investigative procedures related to sound phenomena in gamelan	3	5–7
Researching, Evaluating, and Using Scientific Information for Decision-Making and Action	Analyzing information and determining solutions based on scientific evidence	3	8–10

Based on the instrument's specifications, stimuli were developed based on the acoustic phenomena of the gamelan, linked to the concept of sound waves. These stimuli were used as an authentic context to assess students' science literacy in solving problems related to scientific phenomena and local culture. An example of the developed instrument design is presented in Table 2.

Table 2. Sample Specifications for Science Literacy Test Items Integrating Local Wisdom on Gamelan

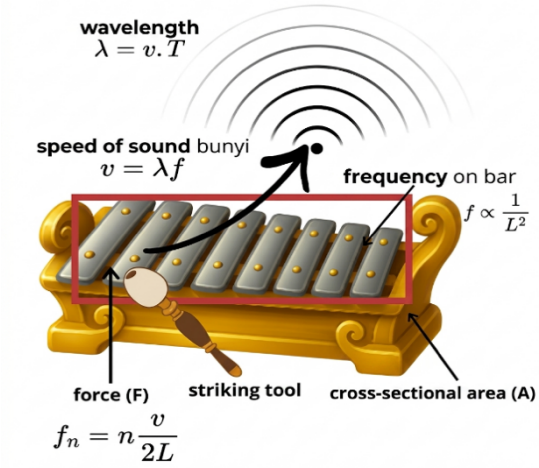

Image of Saron
Source: Personal modification

The Saron is a musical instrument from the Javanese Gamelan that produces sound when its metal bars are struck. The sound is generated by the vibration of the Saron bars, which propagates through the air in the form of longitudinal mechanical sound waves. Each Saron bar has a different length and weight, resulting in varying frequencies. Shorter bars produce higher pitches, while longer bars produce lower pitches. It is this variation in the length of the bars that determines the pitch of the Saron. The strength and weakness of the Saron's sound are determined by the amplitude of the vibration, which depends on the force of the strike. Additionally, the resonator box on the Saron serves to amplify the sound through resonance. Saron players can stop the vibration of a bar with their hands to stop the sound prematurely, which relates to the damping of sound waves.

Science Literacy Competencies	Science Literacy Indicators	Question Indicators	Question
Developing and evaluating designs for scientific investigations and critically interpreting data and evidence	Proposing appropriate experimental designs	Given a reading about the Saron musical instrument related to the concepts of vibration and sound waves, students can propose an experimental design to investigate the relationship between the force of the strike and the loudness of the sound by explaining the experimental steps, independent variables, dependent variables, control variables, and the	Based on the reading, a student wants to determine the effect of striking force on the loudness of the sound produced. Propose a simple experimental design that can be conducted to demonstrate the relationship between striking force, vibration amplitude, and sound intensity on a Saron. Explain the necessary equipment and materials, the experimental steps, and what data should be observed!

	data that needs to be collected scientifically.

Each test item is accompanied by a scoring rubric developed based on science literacy indicators to ensure consistency and objectivity in the assessment of students' responses. The scoring rubric is used during the instrument validation and pilot testing phases. After the design phase was completed, the draft instrument was evaluated through expert validation to assess the appropriateness of the content, construction, language, integration of local wisdom, and scoring rubric. The results of the evaluation by three validators are presented in Table 3.

Table 3. Instrument Validation Results by Validators

Aspect	Validator			Average	Average (%)
	I	II	III		
Content Validity	0.80	0.75	0.90	0.81	81
Construct Validity	0.90	0.85	0.85	0.86	86
Language Validity	1.00	0.95	0.90	0.91	91
Integration of Local Wisdom in Gamelan	0.90	0.90	0.90	0.90	90
Validity of the Assessment Rubric	0.85	0.80	0.80	0.81	81
Final Score				**0.86**	**86**

Based on Table 3, the validation results show that the instrument achieved an average validity percentage of 86%, which falls within the "highly valid" category. The Percentage of Agreement (PoA) between validators also exceeded the minimum threshold of 75%; therefore, the instrument was deemed suitable for limited pilot testing after minor revisions were made in accordance with the experts' recommendations. Next, a pilot study was conducted with 35 11th-grade high school students to determine the empirical validity and reliability of the instrument. The response data were then analyzed using the Rasch model with the assistance of Winsteps software. A summary of the empirical validity based on item fit statistics is presented in Table 4.

Table 4. Summary of Item Validity

Indicator	Result
Number of Items	10
Valid Items	10
Range of Pt Measure Corr	0.36-0.77
Range of Outfit MNSQ	0.52-1.35

The results of the analysis in Table 4 show that all test items (10 items) meet the criteria for empirical validity, with Point-Measure Correlation values ranging from 0.36

to 0.77 and Outfit Mean Square (MNSQ) values falling within the ideal range of 0.52–1.35. The evaluation of the Rasch model's measurement reliability is then summarized in Table 5.

Table 5. Reliability Results

Aspect	Score	Category
Person Reliability	0.76	Fair
Item Reliability	0.91	Very good
Cronbach Alpha	0.81	Good
Person Separation	1.78	Fair

Table 5 shows that the instrument has excellent measurement consistency, as indicated by an Item Reliability value of 0.91 and a Cronbach's Alpha of 0.81. The range of item difficulty levels also varies logitically. The results of the item difficulty distribution are presented in Table 6.

Table 6. Distribution of Item Difficulty Levels

Category	Number of Items
Easy	1
Medium	7
Hard	2

The distribution of the interaction between students' abilities and item difficulty levels is visualized using a Wright Map in Figure 2.

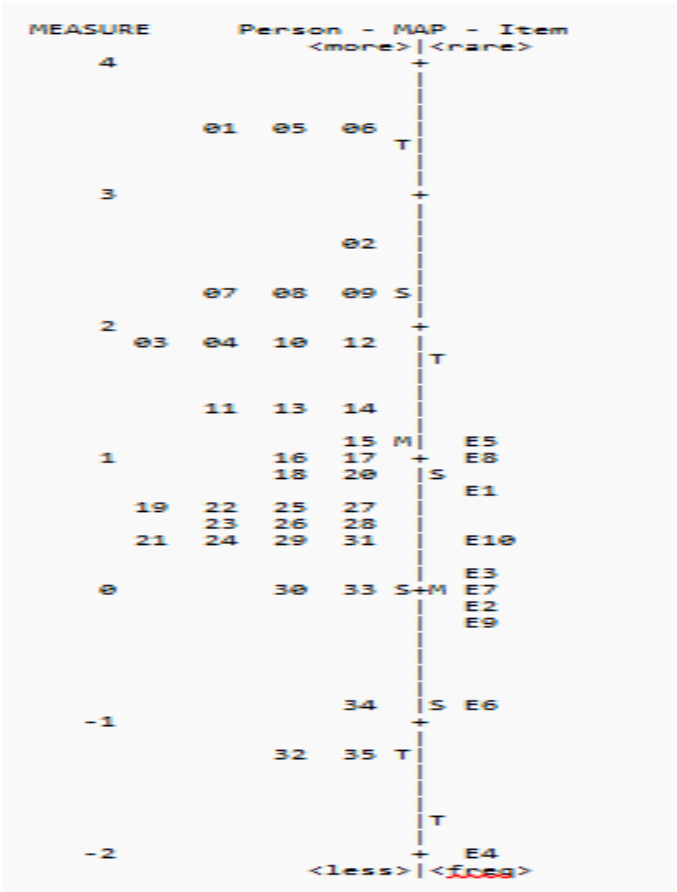

Figure 2. Wright Map Output

Next, the instrument's discriminant power was analyzed based on the separation values and strata generated through Rasch analysis. The results of the discriminant power analysis are presented in Table 7.

Table 7. Particle Size Distribution

Category	Number of Items
Fair	1
Good	8
Very Good	1

The results of the analysis show that all items have discriminant power ranging from moderate to very good, meaning they are capable of distinguishing students based on their level of science literacy.

Discussion

Interest in The Subject Regarding The Decision to Choose a Study Programme

The results of the expert validation indicate that the developed instrument falls into the "highly valid" category, with a high level of agreement among validators. These findings suggest that the aspects of content, construct, language, integration of local gamelan wisdom, and science literacy competencies align with the established measurement objectives. These results align with the research by (Pradana et al., 2022), which states that expert validation is a crucial step in ensuring the content and construct validity of an instrument before it is pilot-tested with students. These results are further supported by Rasch analysis, which indicates that all items meet the item-fit criteria, thereby effectively measuring the same construct: science literacy in the topic of sound waves (Yokhebed et al., 2025). From the perspective of the Rasch Model, the fit of the items to the model indicates that the instrument provides a good representation of the construct being measured (Pradana et al., 2022). Item fit demonstrates that each item functions consistently in measuring the same underlying construct, thereby supporting the validity of the measurement (Long et al., 2023). Consequently, the Rasch Model produces more objective measurements by providing invariant estimates of item difficulty and person ability that are independent of the specific sample used (Tesio et al., 2024). In addition to being valid, the developed instrument also demonstrated good measurement quality based on person reliability, item reliability, Cronbach's Alpha, and the separation index (Khalid et al., 2023). These results indicate that the instrument is capable of producing consistent measurements while distinguishing students based on their ability levels (Natsir et al., 2022). This finding aligns with the research by (Wahyuni & Ragil Widianto Atmojo, 2024), which reported that science literacy instruments with high reliability are capable of providing more accurate and stable information regarding students' abilities. According to Rasch's measurement theory, high reliability indicates that score variation is influenced more by differences in students' abilities than by measurement error (Akbari, 2025). Therefore, the developed instrument not only meets validity criteria but also possesses sufficient measurement precision to be used as a science literacy assessment tool (Kusumaningrat et al., 2026).

The analysis results show that the instrument has a varied distribution of difficulty levels, ranging from easy, moderate, to difficult. This variation is a key characteristic of a well-designed instrument because it allows for the measurement of students' abilities

across various competency levels. This aligns with (CHAN et al., 2023) research, which states that instruments with a varied difficulty range are more effective in comprehensively identifying students' abilities. Items in the difficult category generally require students to integrate the concept of sound waves with gamelan phenomena, evaluate scientific information, and construct arguments based on evidence, whereas easier items tend to measure basic abilities in explaining scientific phenomena. These findings are supported by the results of the discriminative power analysis, which showed a range from adequate to very good, indicating that the instrument can accurately identify differences in science literacy abilities among students. Similar results were also reported by (Iñarrairaegui et al., 2022), who noted that good discriminative power demonstrates the items' ability to consistently distinguish between high- and low-ability students. Thus, the developed instrument possesses adequate item quality to comprehensively measure science literacy skills (Syarifa et al., 2024).

The use of gamelan context in the instrument provides a more contextual assessment experience because students are presented with phenomena closely tied to their cultural environment (Yunior et al., 2022). These findings align with the Context-Based Learning approach, Situated Learning Theory, and Culturally Responsive Pedagogy, which emphasize that conceptual understanding becomes more meaningful when linked to students' social and cultural experiences (Anyichie et al., 2023). Research by (Maftukh Fajar et al., 2023) also indicates that integrating local wisdom into science learning can enhance students' engagement with and understanding of scientific concepts. Therefore, the integration of gamelan serves not only as the context for the questions but also as a means to connect the concept of sound waves to real-life situations (Nasution et al., 2023).

In addition to serving as an evaluation tool, the instrument developed has the potential to be used as a diagnostic assessment in physics education (Dharma et al., 2024). The essay question format allows students to explain scientific phenomena, evaluate information, and use evidence to support decision-making, thereby fostering the development of higher-order thinking skills, which are an essential component of science literacy (Kelvin Kurniawan et al., 2024). These findings align with the research by (Sheptian et al., 2024), which states that science literacy-based essay assessments are effective in revealing students' reasoning and argumentation abilities. Teachers can use the information obtained from the assessment results to identify students' science literacy profiles and design learning strategies that are better suited to their needs (Kusumaningtyas et al., 2025). Thus, this instrument not only supports the improvement of assessment quality in physics education but also contributes to the achievement of SDG 4 (Quality Education) by providing a valid, reliable, contextual, and 21st-century competency-oriented assessment tool (Emiliannur et al., 2026).

CONCLUSION

Fundamental Findings: This study successfully developed an essay-based science literacy assessment instrument integrating the local wisdom of gamelan into the topic of

sound waves based on the ADDIE model up to the development stage. The instrument, consisting of ten essay items aligned with the PISA 2025 science literacy framework, demonstrated high content validity (86%), satisfactory empirical validity, good reliability that is item reliability = 0.91 and Cronbach's Alpha = 0.81; varied levels of difficulty, and adequate discriminative power based on Rasch analysis, indicating that it is capable of objectively and consistently measuring students' science literacy. **Implications:** These findings contribute theoretically by enriching the development of culturally contextualized science literacy assessment based on the PISA 2025 framework and the Rasch measurement model, while practically providing physics teachers with a valid and reliable assessment instrument that supports culturally responsive learning, promotes the preservation of local wisdom, and contributes to the achievement of SDG 4 (Quality Education). **Limitations:** However, this study was limited to the development stage of the ADDIE model and involved only 35 Grade XI students from a single educational context, thereby limiting the generalizability of the findings. **Future Research:** Therefore, future studies should continue the implementation and evaluation stages through large-scale trials involving more diverse participants and educational settings, investigate the effectiveness of the instrument in improving students' science literacy, and explore its adaptation to other physics topics and different local cultural contexts to strengthen its applicability and broader educational impact.

REFERENCES

Abdul Raof, S., & Musta'amal, A. H. (2021). Validity and reliability of instrument competencies framework of agriculture educators: The Rasch measurement model. *International Journal of Educational Methodology*, *7*(3), 411–420. https://doi.org/10.12973/IJEM.7.3.411

Adao, J. M., & Morong, E. (2025). Practices in integrating higher-order thinking skills in english 7 quarterly assessment. *Psychology and Education: A Multidisciplinary Journal*, *41*(4), 476–493. https://doi.org/10.70838/pemj.410406

Adipat, S., & Chotikapanich, R. (2022). Sustainable development goal 4: An education goal to achieve equitable quality education. *Academic Journal of Interdisciplinary Studies*, *11*(6), 174–183. https://doi.org/10.36941/ajis-2022-0159

Akbari, A. (2025). The rasch analysis of item response theory: An untouched area in evaluating student academic translations. *SKASE Journal of Translation and Interpretation*, *18*(1), 50–77. https://doi.org/10.33542/JTI2025-1-05

Alfarizy, Y., Niswati, M., & Jauhariyah, R. (2025). *Analisis Profil Literasi Sains Siswa pada Materi Getaran dan Gelombang dalam Pengembangan Soal PISA Terintegrasi Mitigasi Bencana Gempa dan Tsunami*. *14*(2), 84–90.

Anyichie, A. C., Butler, D. L., Perry, N. E., & Nashon, S. M. (2023). Examining Classroom Contexts in Support of Culturally Diverse Learners' Engagement: An Integration of Self-Regulated Learning and Culturally Responsive Pedagogical Practices. *Frontline Learning Research*, *11*(1), 1–39. https://doi.org/10.14786/flr.v11i1.1115

Bagas Damarsha, A., Khoirun Niza, A., Fitriyah, L., Alan Deta, U., Saputra, O., Studi Pendidikan Fisika, P., Fisika, J., & Matematika dan Ilmu Pengetahuan Alam, F. (2023). Jurnal penelitian pendidikan matematika dan sains: analisis kearifan lokal gamelan (saron) pada konsep fisika gelombang dan bunyi. In *JPPMS* (Vol. 7,

Number 2). https://doi.org/https://doi.org/10.26740/jppms.v7n2.p45-50

Camacho-Sánchez, R., Manzano-León, A., Rodríguez-Ferrer, J. M., Serna, J., & Lavega-Burgués, P. (2023). Game-based learning and gamification in physical education: A systematic review. In *Education Sciences* (Vol. 13, Number 2). MDPI. https://doi.org/10.3390/educsci13020183

CHAN, W. K., ZHANG, L., & OON, E. (2023). Classroom assessment that tailor instruction and direct learning: A validation study. *International Journal of Assessment Tools in Education*, *10*(2), 376–394. https://doi.org/10.21449/ijate.1155679

Dharma, S., -, I., Jaelani, Abd. K., Ariah, F., & -, S. (2024). Desain instrumen pengukuran literasi lingkungan siswa SMA dengan menggunakan pendekatan model rasch. *Jurnal Pendidikan Dan Kebudayaan*, *9*(2), 218–236. https://doi.org/10.24832/jpnk.v9i2.5126

Emiliannur, Novitra, F., Doni, R., Liana, M., Winarno, N., & Hakim, L. (2026). Development and content validation of a HOTS-oriented performance assessment for physics practicum in higher education to support SDG 4. *Jurnal Penelitian Pendidikan IPA*, *12*(3), 681–692. https://doi.org/10.29303/jppipa.v12i3.14471

Fauzan, F., Ansori, R. A. M., Dannur, Moh., Pratama, A., & Hairit, A. (2023). The implementation of the merdeka curriculum (independent curriculum) in strengthening students' character in Indonesia. *Aqlamuna: Journal of Educational Studies*, *1*(1), 136–155. https://doi.org/10.58223/aqlamuna.v1i1.237

Gilbert, R. J. (2006). Competition and Innovation. *Journal of Industrial Organization Education*, *1*(1), 1–23. https://doi.org/10.2202/1935-5041.1007

Heim, A. B., Walsh, C., Esparza, D., Smith, M. K., & Holmes, N. G. (2022). What influences students' abilities to critically evaluate scientific investigations? *PLoS ONE*, *17*(8 August). https://doi.org/10.1371/journal.pone.0273337

Iñarrairaegui, M., Fernández-Ros, N., Lucena, F., Landecho, M. F., García, N., Quiroga, J., & Herrero, J. I. (2022). Evaluation of the quality of multiple-choice questions according to the students' academic level. *BMC Medical Education*, *22*(1). https://doi.org/10.1186/s12909-022-03844-3

Indasa, N., Niswati Rodliyatul Jauhariyah, M., & Phi, J. (2024). Analysis of the science literacy competency profile of high school students on sound wave material. In *Jurnal Phi: Jurnal Pendidikan Fisika dan Fisika Terapan* (Vol. 10, Number 2).

Salsabilla Izzah, I., Jannah, E., & Keguruan dan, F. (2023). Analisis Modul Ajar Berbasis Kurikulum Merdeka. In *Jurnal Literasi dan Pembelajaran Indonesia* (Vol. 3, Number 1).

Kelvin Kurniawan, A., Sandiva, E., Dwi Permata Sari, L., Aulia, Z., & Islam Negeri Fatmawati Sukarno, U. (2024). *Pendampingan pemahaman literasi dan higher order thinking skills (HOTS) dalam konteks pendidikan era digital*. https://ejournal.almarkazibkl.org/index.php/ince

Khalid, N. H. M., Yusof, I. J., Latif, A. A., & Jani, M. D. M. (2023). Psychometric properties of teacher classroom assessment literacy instrument using Rasch model analysis. *International Journal of Evaluation and Research in Education*, *12*(2), 638–646. https://doi.org/10.11591/ijere.v12i2.23607

Kusumaningrat, Nuriyani, E., Afadil, & Nurjannah. (2026). Pengembangan instrumen asesmen ilmu pengetahuan alam sosial berbasis model pisa untuk mengukur kemampuan literasi sains. *Jurnal Ilmiah Profesi Guru (JIPG)*, *7*(1), 66–76. https://doi.org/10.30738/jipg.vol7.no1.a18815

Kusumaningtyas, D. I., Suprapto, N., & Suryanti. (2025). Fostering science literacy: A

comprehensive systematic review of learning models. *Multidisciplinary Reviews*, *8*(9). https://doi.org/10.31893/multirev.2025292

Long, C., Zheng, L., Liu, R., & Duan, Z. (2023). Structural validation and measurement invariance testing of the chinese version of the health literacy scale among undergraduates: Cross-sectional study. *Journal of Medical Internet Research*, *25*(1). https://doi.org/10.2196/48838

Maftukh Fajar, D., Mardiana, A., & Hasanah, R. (2023). *Development of learning videos on vibration, waves, and sounds material at junior high schools based on local wisdom of gamelan Lumajang*. *8*(1). http://jurnalkonstan.ac.id/index.php/jurnal

Mathis, C., Daane, A. R., Rodriguez, B., Hernandez, J., & Huynh, T. (2023). How instructors can view knowledge to implement culturally relevant pedagogy. *Physical Review Physics Education Research*, *19*(1). https://doi.org/10.1103/PhysRevPhysEducRes.19.010105

Maulida, F., & Sunarti, T. (2022). *Pengembangan instrumen tes literasi sains berbasis kearifan lokal di kabupaten Lamongan*. *8*(1).

Nasution Nabilla Nur Almaida, Ardhia Pramesti, C. E. S. A. N., & Suliyanah, D. A. U. S. O. (2023). Jurnal penelitian pendidikan matematika dan sains: Penciptaan konsep fisika pada kearifan lokal kesenian tongklek di Indonesia. In *JPPMS* (Vol. 7, Number 2). https://doi.org/https://doi.org/10.26740/jppms.v7n2.p64-68

Natsir, I., Mayasari, D., & Munfarikhatin, A. (2022). Rasch model: Quality of test instruments to measure students' ability in algebra. *SHS Web of Conferences*, *149*, 01033. https://doi.org/10.1051/shsconf/202214901033

Pradana, M., Silvianita, A., Syarifuddin, S., & Renaldi, R. (2022). The implication of digital organisational culture on firm performance. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.840699

Sakti, S. A., Endraswara, S., & Rohman, A. (2024). Revitalizing local wisdom within character education through ethnopedagogy apporach: A case study on a preschool in Yogyakarta. *Heliyon*, *10*(10). https://doi.org/10.1016/j.heliyon.2024.e31370

Santoso, B., Triono, M., & Zulkifli, &. (2023). Tantangan pendidikan islam menuju era society 5.0: Urgensi pengembangan berpikir kritis dalam pembelajaran PAI di sekolah dasar. *Jurnal Papeda*, *5*(1).

Saputra, A., Sambiri, U., & Fajrhi, N. (2025). Exploring ethnophysics in javanese gamelan: Bridging cultural heritage and STEAM-based physics learning. *SAINTIFIK@: Jurnal Pendidikan MIPA*, *10*(2), 63–74. https://doi.org/10.33387/saintifik.v10i2.11018

Sheptian, R., Ocy, D. R., Triana, D. D., Dwi, L., & Wahyuni, U. (2024). *JISAE (Journal of Indonesian Student Assessment and Evaluation) Developing a validated essay-based assessment instrument to measure science literacy in energy topics*. *10*(2). https://doi.org/10.21009/JISAE

Shufa, N. K. F. (2018). Pembelajaran berbasis kearifan lokal di sekolah dasar: Sebuah kerangka konseptual. *INOPENDAS: Jurnal Ilmiah Kependidikan*, *1*(1). https://doi.org/10.24176/jino.v1i1.2316

Sumintono, B. , & Widhiastrini, W. (2015). *Aplikasi model rasch untuk analisis uji coba instrumen sekolah. Trim Komunikata.*

Syarifa A Alatas, Mohamad Jahja, & Muhammad Yusuf. (2024). Assessment of students' science literacy skills using diagnostic items on the concept of energy with Rasch modeling. *World Journal of Advanced Research and Reviews*, *21*(1), 1356–1362. https://doi.org/10.30574/wjarr.2024.21.1.2697

Tesio, L., Caronni, A., Kumbhare, D., & Scarano, S. (2024). Interpreting results from

Rasch analysis 1. The "most likely" measures coming from the model. *Disability and Rehabilitation*, *46*(3), 591–603. https://doi.org/10.1080/09638288.2023.2169771

Trisiani, D., Maskhanah, S., Lestari, A. B., & Dewi, N. R. (2023). *Peningkatan kemampuan literasi sains peserta didik melalui pendekatan CRT (Culturally Responsive Teaching) dengan alat evaluasi berbasis kearifan lokal kesenian gamelan pada materi getaran dan gelombang*.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press. https://www.hup.harvard.edu/books/9780674576292

Wahyuni, L., & Ragil Widianto Atmojo, I. (2024). *International journal of current science research and review measurement of science literacy skills of elementary school teacher education students: Development and validity testing of assessment instruments*. https://doi.org/10.47191/ijcsrr/V7-i11-27

Yokhebed, Karmadi, R. M. D., Silaen, J., Meidina, S., & Nikmah, N. A. (2025). Analysis of the quality of students' critical reasoning instruments using rasch Model. *Jurnal Penelitian Pendidikan IPA*, *11*(3), 54–60. https://doi.org/10.29303/jppipa.v11i3.9584

Yunior Erlangga, S., Susanti, A., & Fitri, A. (2022). *Pengembangan e-modul fisika materi gelombang dan bunyi berbasis local wisdom alat musik gamelan pada mata kuliah fisika dasar*. *9*(1). https://doi.org/https://doi.org/10.30738/cjipf.v9i2.14154

***Ayunda Zahra Aulia (Corresponding Author)**
Afilliation: Universitas Negeri Surabaya
Address: Jl. Ketintang, Ketintang, Gayungan, Surabaya, Indonesia
Email: raafika.23020@mhs.unesa.ac.id

Titin Sunarti
Afilliation: Universitas Negeri Surabaya
Address: Jl. Ketintang Unesa Campus, 60231, Surabaya, Indonesia
Email : titinsunarti@unesa.ac.id

Joshi Maharani Wibowo
Afilliation: Doctoral Researcher
University of Helsinki, Finland
Address: Helsinki, Uusimaa, Finland
Email: joshi.wibowo@helsinki.fi