JIPF (JURNAL ILMU PENDIDIKAN FISIKA)
p-ISSN: 2477-5959 | e-ISSN: 2477-8451 *Vol. 10 No. 2, May 2025, Page 263-272*

Scientific Literacy Assessment for Physics Instruction by Integration Local Wisdom of *Mandar* Tribe in West Sulawesi, Indonesia

Dewi Sartika [1*)], Nurlina [2], Mutmainna [3]
Universitas Sulawesi Barat, Indonesia[1,2,3]
[*)]Corresponding E-mail: dewi.sartika@unsulbar.ac.id

Received: April 14th, 2025. Revised: May 16th, 2025. Accepted: May 17th, 2025

Keywords :
Scientific Literacy; Physics Instruction; Local Wisdom; Assessment

ABSTRACT
Scientific literacy is an important skill that enables students to apply scientific knowledge to solve daily problems. However, Indonesian students' scientific literacy levels are still relatively low compared to international standards. One potential strategy to address this issue is integrating local wisdom into science learning. This study aims to develop a scientific literacy assessment instrument based on local wisdom that meets the eligibility criteria to improve students' scientific literacy skills comprehensively. Using a research and development (R&D) methodology, this study was conducted in three main stages: product design, product trial, and product assembly. The instrument underwent expert validation and was tested on a limited and larger scale to evaluate readability, item difficulty, item discrimination, and reliability. The findings indicate that the instrument is valid, reliable, and effective for assessing students' scientific literacy. It helps educators diagnose students' abilities to communicate, associate, and apply scientific concepts in real-life and culturally relevant contexts. By connecting science learning with students' cultural backgrounds, this study contributes to enhancing the quality of physics education and improving scientific literacy outcomes among Indonesian students..

INTRODUCTION

Scientific literacy refers to the ability to apply scientific knowledge to solve problems encountered in daily life. It has been a focal point of attention worldwide for many years, as it is considered a key indicator of the success of science education and the quality of human resources in a country [1], [2], [3], [4], [5], [6]. Scientific literacy helps the younger generation face the challenges of the 21st century, which emphasizes critical thinking, analytical skills, creativity, innovation skills, communication, and collaboration [7], [8]. The integration of local wisdom as a learning resource in physics teaching is considered relevant to increasing scientific literacy [9], [10], [11]. The teaching of science, including physics, must be able to equip students with this ability. The Mandar tribe in West Sulawesi, Indonesia, has a rich heritage of local wisdom related to natural phenomena and environmental management. This knowledge can be effectively integrated into physics learning to provide students with a deeper understanding of scientific concepts through culturally relevant examples.

However, the majority of physics teaching methods still focus on teachers and rely on memorization to improve students' ability to answer exam questions [12], [13], [14], [15]. As a result, the average science literacy ability of Indonesian students is far behind international standards. Indonesia's performance in international student assessment programs such as the Program for International Student Assessment (PISA) consistently shows that the scientific literacy of Indonesian students is relatively low compared to other participating countries. In the 2018 PISA measurement, Indonesia was ranked 70 out of 78 countries in science ability [16], [17]. Most students are only able to recognize basic facts but are not able to communicate and relate this knowledge to various science topics, especially in its application in daily life. The low achievement of science literacy of Indonesian students shows that the teaching methods currently applied have not been able to develop students' critical thinking, problem-solving, and scientific reasoning skills optimally.

In addition, students are not trained in solving PISA questions. It is caused by limited teaching materials that encourage higher-level thinking skills, as well as a lack of PISA-like questions. So far, evaluation instruments have also focused more on content, not on science literacy, such as the application of science in daily life, problem-solving, and other science process skills [7], [8], [18], [19].

Therefore, there is a need for reform in physics learning to improve students' science literacy skills, not only limited to mastering content but also the ability to communicate, associate, and apply science knowledge in real life. The development of cultural-based scientific literacy measurement instruments can help teachers diagnose the level of scientific literacy of students and help them become more literate in science by integrating local culture into learning assessments. The purpose of this research is to develop a scientific literacy measurement tool based on local culture that meets the eligibility criteria. Thus, it is hoped that it can improve students' scientific literacy skills comprehensively, not only limited to mastering content but also in the application and communication of science knowledge in real-life contexts.

METHOD

This research employed an educational research and development (R&D) approach. The procedure for developing the scientific literacy assessment consists of three main stages, as illustrated in Figure 1. The development process can be described in detail as follows.

Product Designing Stage

At these stages, the purpose of the test was determined, along with the selection of competencies and materials to be tested, the preparation of the test grid, the formulation of questions and scoring guidelines, and the validation of the assessment. The validation was carried out by two validators whose task was to assess the content validity and the feasibility of the assessment. After the validation process, the results were analyzed, and the assessment was revised based on the experts' suggestions.

Trial Product Stage

The trial phase was divided into two stages: a limited-scale trial and a larger-scale trial. The first stage was a limited-scale trial conducted with three students from a senior high school in Indonesia to assess the readability of the questions. The second stage involved a larger-scale trial with 36 students from the same school to determine the item difficulty index (*P*), item discrimination index *(D*), and reliability coefficient.

Product Assembly Stage
After analyzing the test result data, the items were assembled into a unified scientific literacy assessment instrument.

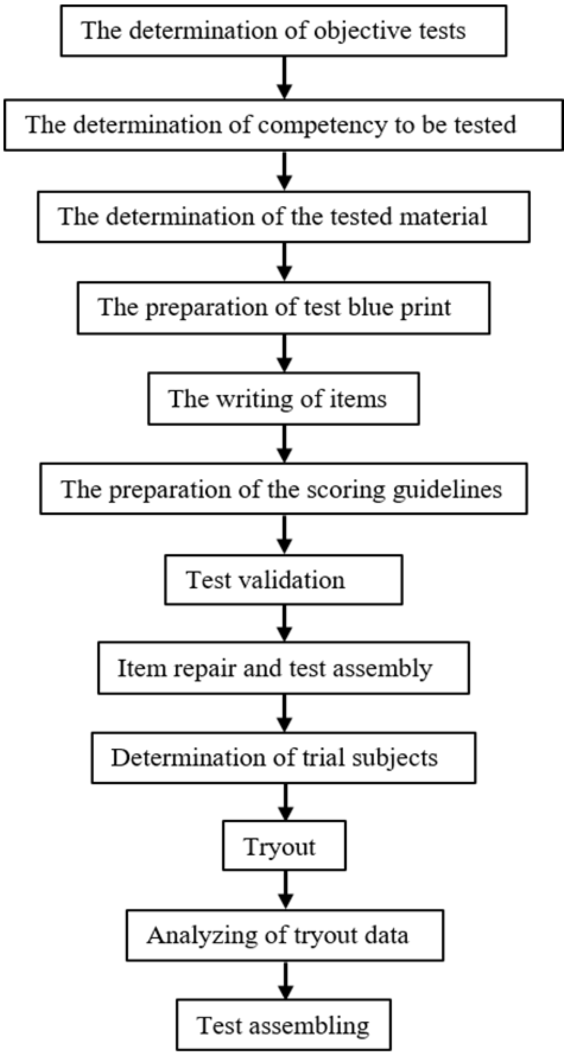


***Fig 1**. Instrument Development Steps*

The participants in this study were divided into two groups. The first group consisted of three students who had participated in the limited trial phase, while the second group consisted of 36 students who participated in the extensive trial phase. Data on instrument feasibility were collected through question item validation and both limited and extensive trials. The Aiken index equation was used to analyze the feasibility data of the instrument, formulated as follows:

$$V = \frac{\sum s}{n(c-1)} \quad (1)$$

The Validator Agreement Index (V) represents the level of agreement among validators in assessing an instrument. The score assigned by each validator (s) is obtained by subtracting the lowest score in the selected category (l_0) from the validator's chosen category score (r), following the formula $s = l_0 - r$. The total number of validators involved in the assessment is denoted by n, while c represents the number of categories available for selection by the validators.

Based on the index calculation results, an instrument can be categorized as valid according to the criteria presented in Table 1.

Table 1. Instrument Validity Criteria

Index Value (*V*)	**Valid Category**
>0.8	Very Valid
0,4 – 0,8	Retained (Needs Revision)
<0.4	Not Valid

Furthermore, the question Item Difficulty Index (P) is determined using the following formula:

$$P = \frac{R}{T} x\ 100\% \tag{2}$$

Where R represents the number of respondents who answered correctly, and T denotes the total number of respondents. The criteria for interpreting the item difficulty index are based on the following guidelines:

Table 2. Item Difficulty Index Criteria

IDI Value	**Difficulty Category**
0,00 – 0.20	Very Difficult
0,21 – 0.40	Difficult
0,41 – 0.60	Moderate (Retained)
0,61 – 0.80	Easy
0,81 – 1.00	Very Easy

The discrimination index (*D*) measures an item's ability to differentiate between high and low-ability students [20], calculated as:

$$D = \frac{High\ (27\%) - Low(27\%)}{N} \tag{3}$$

The criteria for the Item Discrimination Index are presented in Table 3:

Table 3. Item Discrimination Index Criteria

Index Value	**Discriminating Power Category**
-1,00 – 0.20	Very Low
0,21 – 0.40	Low
0,41 – 0.60	Moderate (Retained)
0,61 – 0.80	High
0,81 – 1.00	Very High

Furthermore, the Reliability Coefficient was calculated using the KR-20 method [21].

$$R_{xy} = \frac{n}{n-1}\left[1 - \frac{\sum pq}{S_x^2}\right] \tag{4}$$

Where n represents the total number of test items, p is the proportion of respondents who answered correctly, q is the proportion of respondents who answered incorrectly, and S_x^2 denotes the overall variance of the test.

The reliability coefficient criteria are presented in Table 4. Test instruments with reliability coefficients categorized as moderated, high, or very high are generally considered acceptable as relatively standard testing tools. The detailed criteria are as follows:

Table 4. Reliability Coefficient Criteria

Value	Reliability Coefficient Criteria
0.00 – 0.20	Very Low
0.21 – 0.40	Low
0.41 – 0.60	Moderate (Retained)
0.61 – 0.80	High
0.81 – 1.00	Very High

RESULTS AND DISCUSSIONS

Product Designing

The first step in developing the assessment process was determining the test objectives and identify the materials and competencies to be assessed. The primary objective of the test was to measure students' scientific literacy skills based on three key indicators: (1) explaining phenomena scientifically, (2) evaluating and designing scientific inquiry, and (3) interpreting data and evidence scientifically. Subsequently, the relevant materials and competencies were identified. The selected material focused on the concept of heat, with specific competencies and achievement indicators outlined in Table 5.

Table 5. Basic Competence and Indicators of Competency Achievement

Basic Competence	Indicators of Competency Achievement
Analyzing the effect of heat and heat transfer, which includes the thermal characteristics of a material, capacity, and heat conductivity in everyday life	1. Analyzing the effect of heat on the type of material (conductivity) through observing local cases in everyday life
	2. Analyzing the effect of heat on changes in object temperature (specific heat and heat capacity) through observing local cases in daily life
	3. Analyzing the effect of changes in object temperature on object size (expansion) through life
	4. Analyzing heat transfer by conduction, convection, and radiation through observing local cases in everyday life

The test grid was designed based on scientific literacy indicators and predetermined learning achievement indicators. According to this assessment framework, each scientific literacy indicator was represented by 5 to 9 test items. The development of question items was aligned with the established test grid to ensure consistency and validity. A total of 23 multiple-choice questions were created, each incorporating contextual elements by integrating aspects of the local wisdom of the Mandar tribe.

One example of a question developed for the scientific literacy assessment, which integrates the local wisdom of the Mandar tribe, is presented in Figure 2. As illustrated in the question in Figure 2, the assessment utilized multiple-choice questions. Consequently, the scoring criteria for this prototype followed standard multiple-choice scoring guidelines. A correct answer was awarded 1 point, while an incorrect answer received 0 points. The next stage in the process was assessment validation, which was conducted by two expert validators specializing in physics education research.

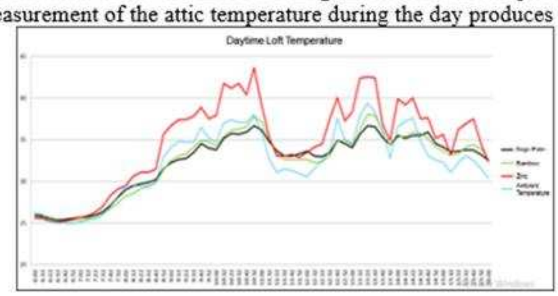
Based on the results of expert validation of the instrument, the content validity coefficient, with a value of 1, meets the criteria for being very valid. An instrument that meets the criteria is very valid, meaning that the instrument is accurate and relevant in measuring scientific literacy skills. The development of this instrument addresses the issue of the lack of scientific literacy assessment tools capable of accurately measuring the literacy levels of local students [18], [19]. However, there are still questions that need to be revised according to the suggestions and inputs from the two expert validators. At this stage, there are nine revised questions, and 3 of them were decided not to be used (*dropout*). It is known that the majority of the revised questions are questions with redaction of core questions and answer choices that are considered less literate by validators, so improvements are made to the redaction of questions and answer choices. In addition, there are also 3 (three) images on 3 (three) different questions, which the validator suggested to be replaced with more contextual images. The revision of the question statement and the replacement of the picture make the instrument more contextual and meet the elements of scientific literacy assessment. Furthermore, the question items in this instrument also contain local cultural elements that make it more unique and more suitable for use in the West Sulawesi region.

The hot air temperature makes people feel uncomfortable sometimes. In daily activities, humans need thermal comfort. Thermal comfort concerns a comfortable room temperature condition that can help the human body maintain its average body temperature of around $37^{0}C$. One of them tries to create a cozy room by choosing the type of roofing material. The roof is one component of a building. Currently, there are various types of roofing materials in the marketplace. One of the indicators in choosing a roofing material is the ability to hold, absorb, and even reflect heat since one of the main functions of the roof is to protect the building from excessive solar heat.

To determine the effect of the thermal characteristics on the comfort of the room, a researcher tried to make a building prototype with roofing materials commonly used by the *Mandar* people. The material is made of zinc roof, sago palm, and bamboo. Temperature measurements are carried out in the morning to noon (when there is sunlight) and at night for 24 hours every 10 minutes.

The prototype with materials: zinc, sago palm, and bamboo

Field testing examined the heat effect of various roofing materials with temperature sensors installed in different positions. Measurement of the attic temperature during the day produces the following data.



Based on the picture, the highest attic temperature is made of zinc during the day. This result is caused by …

a. Zinc roofs are poor conductors compared to sago palm and bamboo roofs.
b. Zinc roofs are good insulators compared to sago palm and bamboo roofs.
c. Zinc roofs quickly transfer the sun's heat compared to sago palm and bamboo roofs.
d. Zinc roofs insulate sun heat more than sago palm and bamboo roofs.
e. Zinc roofs are quickly and widely used in the surrounding environment

***Fig 2**. One of the scientific literacy assessment items that has been developed*

Trial Product

Furthermore, the results of the limited trial demonstrated that the readability of the questions, including both the wording and the accompanying images, was categorized as good. As a result, no further revisions were necessary. This finding indicates that the developed instrument was clear and easily understood by students, ensuring its effectiveness in assessing scientific literacy.

Nevertheless, the extensive trial revealed several unsatisfactory results regarding the developed instrument. The findings indicated that: (1) the reliability coefficient of the instrument was still relatively low, (2) more than 50% of the items had a poor discrimination index, and (3) 40% of the items fell into the "very easy" category based on the difficulty index. These results suggest that further revisions and refinements are necessary to enhance the instrument's reliability, item discrimination, and overall effectiveness in measuring scientific literacy.

The reliability coefficient, which is classified as a low category, according to [19], is reasonable even though the reliability index in the world of education that is tolerated to be accepted as a relatively standard test tool is a test result that is at a value of 0.40 – 1.00, namely in the medium, high, and very high categories [22], [23], [24]. Therefore, it is highly recommended to revise several questions to improve the quality of the reliability of the instrument that has been developed.

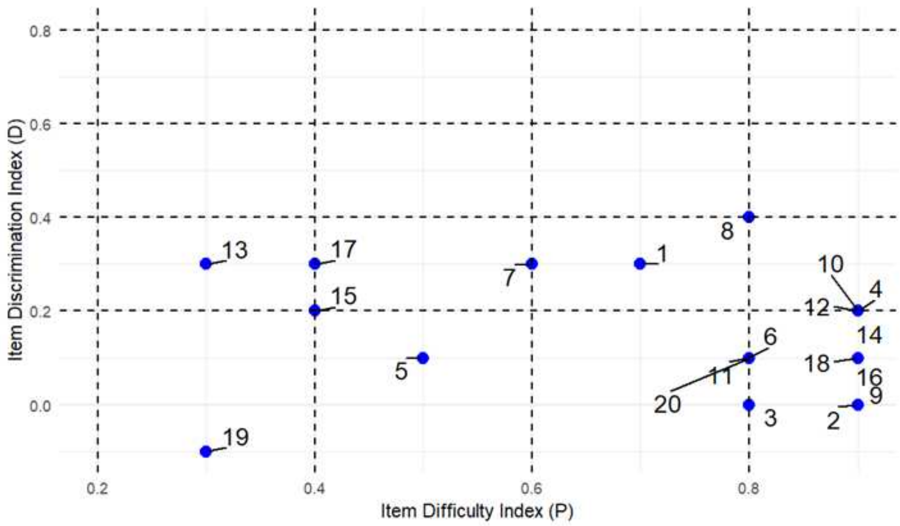


Fig 3. *Item characteristic map*

Furthermore, this revision suggestion is strengthened by the results of the discrimination index and the item difficulty index, as shown in Figure 3. A total of 11 questions in the results of the analysis of the power discrimination index obtained a score below 0.20, two questions obtained a score of 0, and 1 question obtained a value of -1. From the results of the literature review, it is known that if the item discrimination index has a value of zero, this means that the item is not able to distinguish between students who know the correct answer and students who do not know. Some of the factors that can affect the low item discrimination index are (1) the item is too easy or too difficult so that it allows all students to be wrong or all students are right, (2) the item is confusing as a result of ambiguous sentence construction (giving rise to double interpretation), this is because the question sentence is too long. For item difficulty index data, it is known that 8 out of 20 questions are categorized as very easy, with a score of 0.81 – 1.00. It is known that the tolerance limit as a standard test is a question item that has an item difficulty index of 0.30 – 0.70. Question items that are too easy result in almost all students being able to answer correctly on these items, so it will be difficult for researchers to know whether the question items can test students' literacy skills properly. From this explanation, it can be further understood the importance of revising several items of this test instrument that has been developed. In addition, to maintain the quality of students' answers purely from the results of their thoughts, it is recommended that strict supervision is carried out during the test. It is needed to avoid students cheating on each other and produce mostly correct answers. Therefore, it was decided that there were 3 question items to be revised because the item was categorized as lacking for the difference index as well as being categorized as very easy for the difficulty index.

On the other hand, based on Figure 4, which presents the results of the larger test, it is also known that although the average scores for each indicator fall within the high category, students demonstrate greater proficiency in answering scientific literacy questions related to Indicator (1), *Explaining Scientific Phenomena*, compared to Indicator (2), *Evaluating and Designing Scientific Investigations*, and Indicator (3), *Interpreting Scientific Data and Evidence*. However, these findings address the issue of low scientific literacy scores among students. These results align with previous studies [25], [26], [27], [28], [29], [30], [31], [32]. This outcome can be attributed to the fact that the number of questions in Indicator (1) is higher than in the other two indicators. Additionally, students are frequently exposed to conceptual questions and real-life science phenomena in physics lessons, particularly during the stimulus stage. As a result, they are more accustomed to answering physics questions that relate to their daily experiences. However, for Indicator (2), *Evaluating and Designing Scientific Investigations*, and Indicator (3), *Interpreting Data and Evidence*, students still rarely meet and work on problems with this type. Hence, the score on this indicator is not as high as the previous Indicator (1), *Explaining Scientific Phenomena*.

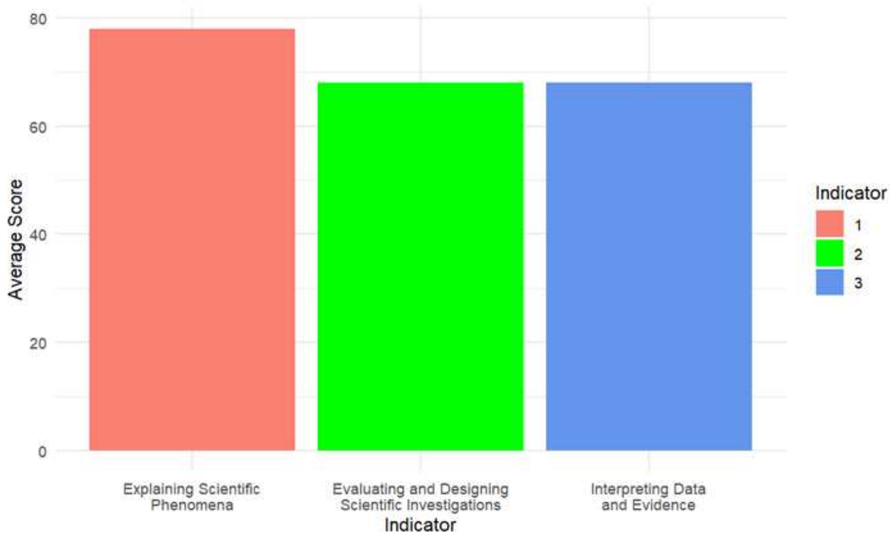


***Fig 4**. Scientific literacy scores per indicator*

Product Assembly

The product assembly stage marked the final phase in developing the scientific literacy assessment instrument. At this stage, revisions were made based on findings from both limited and large trials. Items with poor discrimination and very high ease levels were either revised or eliminated to improve the instrument's psychometric quality. Specifically, three items were revised due to overlapping weaknesses in discrimination indices. This refinement process ensured that only the most representative and functional items were retained in the final instrument. Furthermore, attention was given to content coverage, clarity, and cultural contextualization. As a result, the final instrument is not only valid and reliable but also culturally responsive, making it a practical tool for assessing students' scientific literacy in the physics learning context within Indonesia, particularly in culturally rich regions such as West Sulawesi.

CONCLUSION AND SUGGESTION

The developed scientific literacy assessment instrument effectively measures students' literacy skills across three key indicators, incorporating local wisdom from the Mandar tribe for contextual relevance. Expert validation confirmed high content validity, but revisions were needed to improve question clarity and contextualization. While the limited trial showed good readability, the extensive trial revealed low reliability, poor item discrimination, and many overly easy questions. Students performed best in Explaining Scientific Phenomena, likely due to greater exposure to physics lessons, while they struggled more with Evaluating Scientific Investigations and Interpreting Data. This assessment tool provides a culturally relevant approach to measuring scientific literacy and can support teachers in diagnosing students' literacy levels. Despite its high validity, further improvements are needed to enhance reliability and discriminatory power.

ACKNOWLEDGMENTS

We acknowledge the support from the Universitas Sulawesi Barat for funding this research.

REFERENCES

[1] Ogunkola, B. J. (2013). Scientific literacy: Conceptual overview, importance and strategies for improvement. *Journal of Educational and Social Research*, *3*(1), 265-274.

[2] Kalkan, Ö. K., Altun, A., & Atar, B. (2020). Role of teacher-related factors and educational resources in science literacy: An international perspective. *Studies in Educational Evaluation*, *67*, 100935.

[3] Marmoah, S., Poerwanti, J. I. S., & Suharno. (2022). Literacy culture management of elementary school in Indonesia. *Heliyon*, *8*(4).

[4] Raman, R., Venugopalan, M., & Kamal, A. (2024). Evaluating human resources management literacy: A performance analysis of ChatGPT and bard. *Heliyon*, *10*(5).

[5] Rehman, A., Jingdong, L., & Hussain, I. (2015). The province-wise literacy rate in Pakistan and its impact on the economy. *Pacific Science Review B: Humanities and Social Sciences*, *1*(3), 140-144.

[6] Şentürk, C., & Sari, H. (2018). Investigation of the contribution of differentiated instruction into science literacy. *Qualitative Research in Education*, *7*(2), 197-237.

[7] Istyadji, M. (2023). Conception of scientific literacy in the development of scientific literacy assessment tools: A systematic theoretical review. *Journal of Turkish Science Education*, *20*(2), 281-308.

[8] Noor, M. S. A. M. (2021). Assessing secondary students' scientific literacy: A comparative study of suburban schools in England and Malaysia. *Science Education International*, *32*(4), 343-352.

[9] Hartini, S., Firdausi, S., Misbah, M., & Sulaeman, N. F. (2018). The development of physics teaching materials based on local wisdom to train saraba kawa character. *Jurnal Pendidikan IPA Indonesia*, *7*(2), 130-137.

[10] Hidayati, N. A., Waluyo, H. J., & Winarni, R. (2020). Exploring the Implementation of Local Wisdom-Based Character Education among Indonesian Higher Education Students. *International Journal of Instruction*, *13*(2), 179-198.

[11] Suastra, I. W., Jatmiko, B., Ristiati, N. P., & Yasmini, L. P. B. (2017). Developing characters based on local wisdom of bali in teaching physics in senior high school. *Jurnal Pendidikan IPA Indonesia*, *6*(2), 306-312.

[12] Adams, B., Thomas, L., Moens, M., & Valcke, M. (2023). University teachers as versatile masters: Evaluating the effectiveness of a professional development programme on student-centred teaching competencies. *Studies in Educational Evaluation*, *77*, 101260.

[13] Schiering, D., Sorge, S., Tröbst, S., & Neumann, K. (2023). Course quality in higher education teacher training: What matters for pre-service physics teachers' content knowledge development?. *Studies in Educational Evaluation*, *78*, 101275.

[14] Lehesvuori, S., Lehtinen, A., Hämäläinen, R., Maunuksela, J., & Koskinen, P. (2023). Student-centredness in physics laboratory teaching sessions. *Learning, Culture and Social Interaction*, *43*, 100773.

[15] Balta, N., Japashov, N., Bakytkazy, T., Abdikadyr, B., & Nurgaliyeva, K. (2024). High school physics teachers' pedagogical discontentment: The effect of curriculum improvement and professional development programs. *Heliyon*, *10*(2).

[16] OECD. (2023). *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*. Paris.

[17] Pahrudin, A., Irwandani, I., Triyana, E., Oktarisa, Y., & Anwar, C. (2019). The Analysis of Pre-Service Physics Teachers in Scientific Literacy: Focus on the Competence and Knowledge Aspects. *Jurnal Pendidikan IPA Indonesia*, *8*(1), 52-62.

[18] Coppi, M., Fialho, I., & Cid, M. (2023). Developing a scientific literacy assessment instrument for Portuguese 3rd cycle students. *Education Sciences*, *13*(9), 941.

[19] Zhang, L., Liu, X., & Feng, H. (2023). Development and validation of an instrument for assessing scientific literacy from junior to senior high school. *Disciplinary and Interdisciplinary Science Education Research*, *5*(1), 21.

[20] Adams, W. K., & Wieman, C. E. (2011). Development and validation of instruments to measure learning of expert-like thinking. *International journal of science education*, *33*(9), 1289-1312.

[21] Tuckman, B. W., & Harper, B. E. (2012). *Conducting educational research*. Rowman & Littlefield Publishers.

[22] Lisufiana, D., Supriyanto, T., & Khumaedi, M. (2019). Validity and Reliability Content of Instrument of Assessments Mengalihaksarakan Serat Wulangreh Pupuh Gambuh Class Viii. *Journal of Research and Educational Research Evaluation*, *8*(2), 165-170.

[23] Mohajan, H. K. (2017). Two criteria for good measurements in research: Validity and reliability. *Annals of Spiru Haret University. Economic Series*, *17*(4), 59-82.

[24] Adi, S., Nurharsono, T., Billiandri, B., & Soenyoto, T. (2023, November). Validity and Reliability of Instruments in Physical Education Learning Multimedia. In *Proceedings of International Conference on Physical Education, Health, and Sports* (Vol. 3, pp. 1-13).

[25] Smith, S. E., Mason, J. C., & Bowden, M. (2020). Local wisdom in regenerative teacher practices. *Australian journal of teacher education (online)*, *45*(9), 92-107.

[26] Jamaluddin, A. B., Zubaidah, S., Mahanal, S., & Gofur, A. (2022). Exploration of the Indonesian Makassar-Buginese" Siri'" Educational Values: The Foundation of Character Education. *International Journal of Evaluation and Research in Education*, *11*(1), 10-19.

[27] Hartini, S., Firdausi, S., Misbah, M., & Sulaeman, N. F. (2018). The development of physics teaching materials based on local wisdom to train saraba kawa character. *Jurnal Pendidikan IPA Indonesia*, *7*(2), 130-137.

[28] Hidayati, N. A., Waluyo, H. J., & Winarni, R. (2020). Exploring the Implementation of Local Wisdom-Based Character Education among Indonesian Higher Education Students. *International Journal of Instruction*, *13*(2), 179-198.

[29] Suastra, I. W., Jatmiko, B., Ristiati, N. P., & Yasmini, L. P. B. (2017). Developing characters based on local wisdom of bali in teaching physics in senior high school. *Jurnal Pendidikan IPA Indonesia*, *6*(2), 306-312.

[30] Setiawan, B., Innatesari, D. K., Sabtiawan, W. B., & Sudarmin, S. (2017). The development of local wisdom-based natural science module to improve science literation of students. *Jurnal Pendidikan IPA Indonesia*, *6*(1).

[31] Dwianto, A., Wilujeng, I., Prasetyo, Z. K., & Suryadarma, I. G. (2017). The development of science domain based learning tool which is integrated with local wisdom to improve science process skill and scientific attitude. *Jurnal Pendidikan IPA Indonesia*, *6*(1).

[32] Ramlan, R., Iskandar, D., Permana, J., & Husin, M. R. (2023). Character Values of Elementary School Education from the Perspective of Local Wisdom of Sundanese Culture. *Journal of Educational and Social Research*, *13*.