



Analysis of Item Response Theory in Science Literacy Assessment Containing Local Wisdom and Pancasila Student Profile

Dewi Dewantara*

Physics Education Study Program, Universitas Lambung Mangkurat, South Kalimantan,
Indonesia

Mustika Wati

Physics Education Study Program, Universitas Lambung Mangkurat, South Kalimantan,
Indonesia

Sri Hartini

Physics Education Study Program, Universitas Lambung Mangkurat, South Kalimantan,
Indonesia

Surya Haryandi

Physics Education Study Program, Universitas Lambung Mangkurat, South Kalimantan,
Indonesia

Ismi Fajariah

Physics Education Study Program, Universitas Lambung Mangkurat, South Kalimantan,
Indonesia

Nuzuluna Aulia Syifa

Physics Education Study Program, Universitas Lambung Mangkurat, South Kalimantan,
Indonesia

Muhammad Riduan Fitri

Physics Education Study Program, Universitas Lambung Mangkurat, South Kalimantan,
Indonesia

*Correspondence author: dewantara_pfis@ulm.ac.id

DOI: <http://dx.doi.org/10.18592/tarbiyah.v13i2.13771>

Key Words:

assessment
item response theory
local wisdom
pancasila student profile
science literacy

Abstract:

Science literacy assessment is critical because it helps measure people's understanding of scientific concepts and ensures the effectiveness of science education. This study aimed to identify the quality of the instrument based on item response theory in assessing high school physics science literacy. This instrument is found in particle dynamics material. This type of research is quantitative descriptive. The location of data collection was in three schools in Banjarmasin City. The sample was obtained by cluster random sampling. The validity of the Rasch program was obtained from the results of the fit order items compared to the MNSQ and ZSTD value criteria. The validity test and discriminatory power analysis results showed that three

Received : 13 August 2024

Revised : 17 September 2024

Accepted : 28 November 2024

Published : 24 December 2024

invalid questions had to be discarded. The results showed that the developed science literacy instrument had a level of difficulty in the middle, so the instrument was suitable for both high-ability and low-ability students. This study implies that this assessment can help students relate physics concepts to local culture, increasing the relevance of learning and producing students who are not only academically competent but also have social and cultural awareness.

To cite this article: Dewantara, D., Wati, M., Hartini, S., Haryandi, S., Fajariah, I., Syifa, N.A., Fitri, M.R. (2024). Analysis of Item Response Theory in Science Literacy Assessment Containing Local Wisdom and Pancasila Student Profile. *Tarbiyah: Jurnal Ilmiah Kependidikan*. Vol 13 (2), 229-243.

This is an open access article under the CC-BY License
(<https://creativecommons.org/licenses/by/4.0/>)



Introduction

Having life skills that will help students thrive over time is a learning necessity in the 21st century (Afnan et al., 2023; Dewantara et al., 2019; Syafitri, 2023; Wel Ningsi & Mahaputra Iskandar, 2023). Scientific literacy is one of the life skills that students need today. It is expected that students will have the ability to interpret scientific data and evidence, evaluate and design scientific research, explain scientific phenomena, and interpret data and evidence (Hidayat et al., 2021; Sari et al., 2023a; Sholihah et al., 2023). When physics material is linked to local wisdom, the physics learning process can become meaningful learning for students (Ali & Zaini, 2023; Bramastia et al., 2023; Khusniati et al., 2023; Putra et al., 2020; Rusmansyah et al., 2023; Sari et al., 2023b; Sholihah et al., 2023; Siswanto et al., 2023).

The meaningfulness of learning in achieving 21st-century skills must be based on a strong foundation as the young generation of Indonesia. The profile of Pancasila students is the main foundation for building the character and morality of the young generation of Indonesia. Through a deep understanding of the values of Pancasila, students can internalize basic principles such as justice, unity, democracy, the one and only God, and just and civilized humanity (Dewi et al., 2023; Nuraini et al., 2024; Santika, 2020). The profile of Pancasila students not only helps form individuals with character and responsibility but also strengthens national identity and unity. By having a strong understanding of Pancasila, students can become agents of change who promote peace, tolerance, and justice in society and can face various global challenges with an attitude based on the nation's noble values (N. Huda et al., 2024; Ismail et al., 2021).

Measuring 21st-century skills is very important to determine the level of achievement of learning indicators and to collect information about students' learning development on various topics (Alti et al., 2021; Chasanah et al., 2022). Assessment is needed to determine the level of student ability that still lacks abilities that need to be improved. The assessment's quality determines the measurement's success (Aulia et al., 2021a). Science literacy assessments are important because they help measure people's understanding of scientific concepts, ensure the effectiveness of science education, and enable the development of more effective learning strategies to improve understanding and skills in science (Alifiyah et al., 2020; Aulia et al., 2021b).

In 2023, the proposing team received a DRTPM grant to develop a high school physics science literacy assessment based on Local Wisdom and the Pancasila Student Profile. The assessment developed needs to be tested for quality for further use according to its usefulness. The quality of the assessment is reviewed from the validity, reliability, level of difficulty, and discriminatory power (A. A. S. Huda & Khairina, 2023; M Wati, 2021; Setiawan Putra et al., 2018, 2021; Wati et al., 2015, 2018, 2021; Wati & Mahtari, 2017; Wijayadi et al., 2023). Validity is the accuracy of the test instrument in the form of descriptive questions in measuring students' scientific literacy. Reliability is the extent to which the test instrument created can be trusted, and the results are not affected by circumstances (Fitri, 2007; Suharsono & Istiqomah, 2014; Yusup, 2018). Reliability testing is carried out to measure the level of consistency of the instruments developed (Azwar, 2012; Matondang, 2009).

The level of difficulty is a measure of the ease of a question. Items with low criteria imply that students find it easy to solve the problem, and items with complex

criteria imply that students find it challenging to solve the problem (Purwanto, 2016; E. P. Widoyoko, 2019). The discriminatory power of a test instrument is the ability of a test instrument, in the form of descriptive questions, to differentiate between high and low-achieving students (Mansur, 2008; S. E. P. Widoyoko, 2017). The discriminating power is calculated using the number of correct and incorrect answers of students from the lower and upper groups. The quality of the assessment that has been developed will be tested through Item Response Theory analysis with Rasch modeling using the Winstep application.

The general problem formulation of this study is "How is the quality of the assessment of high school physics science literacy based on Local wisdom and the Pancasila Student Profile?". The specific research questions are: a) Validity of the assessment of high school physics science literacy based on Local wisdom and the Pancasila Student Profile; b) Reliability of the assessment of high school physics science literacy based on Local wisdom and the Pancasila Student Profile; c) Level of difficulty of the assessment of high school physics science literacy based on Local wisdom and the Pancasila Student Profile; d) Distinguishing power of the assessment of high school physics science literacy based on Local wisdom and the Pancasila Student Profile. The use of IRT to evaluate assessment quality offers a modern statistical framework, allowing for a more nuanced understanding of item validity, reliability, difficulty, and distinguishing power.

The advantage of this assessment is that it integrates the Pancasila student profile and local wisdom in each question so that there is cultural relevance and local context in evaluating science literacy. The general purpose of the study is to identify the quality of the Instrument based on item response theory in the assessment of high school physics science literacy based on Local wisdom and the Pancasila Student Profile, which is seen from: a) Validity of the assessment of high school physics science literacy based on Local wisdom and the Pancasila Student Profile; b) Reliability of the assessment of high school physics science literacy based on Local wisdom and the Pancasila Student Profile; c) Level of difficulty of the assessment of high school physics science literacy based on Local wisdom and the Pancasila Student Profile; d) Distinguishing power of the assessment of high school physics science literacy based on Local wisdom and the Pancasila Student Profile

Research Method

This type of research is quantitative descriptive research. The location of data collection is at three schools in Banjarmasin city. The instrument used is an assessment of high school physics science literacy based on Local wisdom and the Pancasila Student Profile. Students selected through cluster random sampling will answer the questions in the assessment. Student answer data is collected and analyzed according to the answer key. The data is entered into Excel and then entered into Winstep to be analyzed using the Rasch Model.

1. Validity

The validity calculation is reviewed from the MNSQ outfit according to table 1 (Sumintono & Widhiarso, 2015). The criteria outlined in Table 1 specify the validity

requirements for an assessment based on outfit mean square (MNSQ) and point measure correlation (Pt Measure Corr).

Table 1. Validity Requirements

<i>Output</i>	Kriteria
$0,5 < outfit\ MNSQ < 1,5$ $Pt\ Measure\ Corr > 0$	Valid
Selain kriteria di atas	Tidak Valid

Outfit MNSQ ($0.5 < MNSQ < 1.5$) indicates that the items fit the Rasch model appropriately. Values below 0.5 may suggest items are too predictable, while values above 1.5 indicate unexpected responses. Staying within this range implies that items are balanced in difficulty and contribute meaningfully to the assessment. Point Measure Correlation ($Pt\ Measure\ Corr > 0$) means that students' ability increases, and their likelihood of correctly answering items. This relationship confirms that the items function as intended, aligning well with the student's capabilities. Together, these criteria ensure the validity of the assessment; items meeting these standards are considered valid, while those that do not are classified as invalid.

2. Reliability

The reliability value can be seen in the output table: summary of measure items with In the Rasch model, to determine the reliability value of the question, it can be analyzed through the Item Reliability and Person Reliability values based on the criteria in Table 2. (Sumintono & Widhiarso, 2015).

Table 2. Reliability Criteria in Rasch Modeling

Reliability (Person/Item)	Interpretation
>0.94	Excellent
0.91 - 0.94	Very Good
0.81 - 0.90	Good
0.67 - 0.80	Fair
<0.67	Weak

3. The level of difficulty

In Rasch modeling, the difficulty level of the question items is categorized based on the Measure logit and the Standard Deviation (SD) value of the item logit, which are divided into four categories (Sumintono & Widhiarso, 2015) as follows.

Table 3. Criteria for the Level of Difficulty of Question Items with Rasch Modeling

Measure (logit)	Interpretation of Item Difficulty
$Measure\ logit < -SD\ logit$	Very easy items
$-SD\ logit \leq Measure\ logit \leq 0$	Easy items
$0 \leq Measure\ logit \leq SD\ logit$	Difficult items
$Measure\ logit > SD\ logit$	Very difficult items

In addition, the level of difficulty of the test items can also be seen through the item map output.

4. Differentiating power

Unlike the classical theory approach, in Rasch modeling, individual ability level analysis is used as a differentiating power to distinguish the abilities of students who can answer questions and are unable to answer questions. In addition, it can also be used to identify respondent groups based on the respondent separation index. The greater the item separation value, the better the quality of the instrument in terms of overall respondents and test items because it can identify respondent groups and item groups (Sumintono & Widhiarso, 2015).

Another equation to find out the grouping more precisely is the strata equation (H):

$$H = \frac{[(4 \times SEPARATION) + 1]}{3} \quad (2)$$

Based on the item separation index, students with high, medium, and low abilities can be determined.

Results and Discussion

The study was conducted in 3 state senior high schools in Banjarmasin City, which is the population in this study. The total sample of senior high school students in this study was 94 people. The study was conducted to describe the quality of the assessment of high school Physics science literacy based on local wisdom and the Pancasila Student Profile in terms of validity, reliability, level of difficulty, and discriminatory power analyzed using Winstep based on the Rasch model.

1. Validity Result

Validity in the Rasch program is obtained from the results of the item fit order compared to the MNSQ and ZSTD value criteria. Questions are said to be valid if they meet the following criteria.

- a) The Outfit MNSQ (Mean Square) value received is $0.5 < \text{Outfit - MNSQ} < 1.5$
- b) The Outfit ZSTD (Z - Standard) value received is $-2.0 < \text{ZSTD} < +2.0$
- c) Pt Measure Corr (Point Measure Correlation) value: $0.4 < \text{Point Measure Corr} < 0.85$ (Suminto & Widhiarso, 2015)

The summary of the item validity statistics can be seen in Table 4. below.

Table 4. Item Validity

Science Literacy Indicators	Item
Explaining scientific phenomena	G1, G2, G5, G8, G9, G18
Interpreting scientific data and evidence	H3, H4, H6, H11, H12, H14, H15, H16, H17
Evaluating and designing scientific investigations	I7, I10, I13, I19

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT		OUTFIT		PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	Item
					MNSQ	ZSTD	MNSQ	ZSTD					
12	7	94	1.69	.40	1.13	.51	2.21	2.14	-.18	.16	92.6	92.5	N12
3	18	94	.55	.27	1.13	.84	1.74	1.62	-.07	.25	80.9	80.8	N3
11	9	94	1.40	.36	1.12	.53	1.62	1.47	-.07	.18	90.4	90.4	N11
14	33	94	-.35	.23	1.02	.28	1.22	1.67	.24	.31	68.1	67.3	N14
5	26	94	.03	.24	1.04	.40	1.10	.67	.22	.29	70.2	72.9	N5
18	22	94	.28	.25	1.09	.71	1.07	.41	.17	.27	74.5	76.7	N18
9	16	94	.71	.28	.97	-.13	1.08	.38	.24	.24	83.0	83.0	N9
8	30	94	-.19	.23	1.07	.77	1.06	.49	.22	.30	67.0	69.6	N8
17	20	94	.41	.26	1.06	.45	.99	.02	.20	.26	78.7	78.7	N17
15	9	94	1.40	.36	1.04	.22	.95	-.01	.15	.18	90.4	90.4	N15
16	46	94	-.99	.22	1.04	.56	1.02	.23	.30	.34	59.6	64.2	N16
10	42	94	-.80	.22	1.00	-.04	1.00	-.02	.33	.33	66.0	64.5	N10
2	33	94	-.35	.23	.99	-.05	.92	-.59	.34	.31	63.8	67.3	N2
7	41	94	-.75	.22	.94	-.85	.92	-.83	.40	.33	71.3	64.6	N7
4	18	94	.55	.27	.88	-.75	.74	-1.13	.42	.25	80.9	80.8	N4
1	68	94	-2.12	.24	.87	-1.17	.82	-1.19	.49	.33	77.7	74.1	N1
13	42	94	-.80	.22	.86	-2.18	.84	-1.84	.51	.33	78.7	64.5	N13
6	15	94	.79	.29	.85	-.84	.68	-1.23	.44	.23	84.0	84.0	N6
19	56	94	-1.48	.22	.85	-1.93	.81	-2.09	.53	.34	73.4	66.7	N19
MEAN	29.0	94.0	.00	.26	1.00	-.14	1.09	.07			76.4	75.4	
P.SD	16.4	.0	.99	.05	.10	.88	.37	1.27			9.2	9.4	

Figure 1. Measure Order results with Rasch modeling

Based on the results of Table 1, there were three invalid questions, namely questions 3, 11, and 12, with indicators interpreting data and scientific evidence and questions with code H. In question number 3, the MNSQ outfit value was obtained at 1.74, so it did not meet the MNSQ outfit criteria. Furthermore, in question number 11, the MNSQ outfit value was obtained at 1.62, and the Pt Measure Corr was -0.07, which did not meet the criteria; this question item was disqualified. Question number 12 also obtained an MNSQ outfit value of 2.21 and a Pt Measure Corr of -0.18, which did not meet the criteria. The three items that did not meet the validity criteria were disqualified, so of the 19 questions analyzed, 16 valid questions met the MNSQ Outfit value, ZSTD Outfit value, and Point Measure Correlation value. Questions that are declared valid mean that there is a parallel between the question score and the total score, thus indicating that the question can measure the student's abilities, whereas if the question score cannot provide support for the total score, the question is declared invalid. Test instruments with invalid measuring power cannot provide information regarding student abilities (Chan et al., 2013; Sumintono & Widhiarso, 2015).

2. Reliability Item

The reliability of an instrument is the extent to which measurements taken repeatedly will produce the same information. A test is reliable if there is consistency between one test item and another. The instrument's reliability value can be seen through statistical analysis in the Summary Statistics menu in Figure 2 below.

SUMMARY OF 19 MEASURED Item

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	29.0	94.0	.00	.26	1.00	-.14	1.09	.07
SEM	3.9	.0	.23	.01	.02	.21	.09	.30
P.SD	16.4	.0	.99	.05	.10	.88	.37	1.27
S.SD	16.8	.0	1.01	.05	.10	.90	.38	1.30
MAX.	68.0	94.0	1.69	.40	1.13	.84	2.21	2.67
MIN.	7.0	94.0	-2.12	.22	.85	-2.18	.68	-2.09
REAL RMSE	.28	TRUE SD	.95	SEPARATION	3.43	Item	RELIABILITY	.92
MODEL RMSE	.27	TRUE SD	.95	SEPARATION	3.53	Item	RELIABILITY	.93
S.E. OF Item MEAN = .23								

Figure 2. Summary Statistic Results with Rasch Modeling

Based on statistical analysis data, the Item Reliability value is 0.92, which means that the quality of the instrument is in the outstanding category. This shows that this instrument is reliable and is included in the very high criteria for its reliability. The higher the reliability value, the higher the consistency value, and a test with high-reliability consistency is accurate. So, the developed science literacy test instrument has good consistency.

The instrument's reliability calculation results also show the separation value used to determine the grouping of students. By using the equation $H = \{(4 \times \text{Separation}) + 1\} / \text{three}$ and a separation person value of 3.43, a value of 4.90 is obtained (rounded to 5), which means that this instrument groups students into five groups, namely very high, high, sufficient, low, and very low. The greater the separation value, the better the overall quality of the instrument because it can identify groups of respondents and items.

3. The level of difficulty

The level of difficulty measures the ease of the question. If the question has a low criterion, it is easy to agree on or solve with students, while questions with a problematic criterion are challenging to agree on or work on. The results of the difficulty of the question item are presented in Table 5 below.

Table 5. Results of the Distribution of the Level of Difficulty of Question Items with Rasch Modeling

Measure (logit)	Interpretation of Item Difficulty	Number of Item
<i>Measure logit < - 0,99</i>	Very easy items	2
<i>- 0,99 ≤ Measure logit ≤ 0,00</i>	Easy items	7
<i>0,00 ≤ Measure logit ≤ 0,99</i>	Difficult items	7
<i>Measure logit > 0,99</i>	Very difficult items	3

Based on Table 5. it is obtained that the questions with the highest difficulty level are three questions with a logit value >0.99 , so these three questions are included in the difficult category. The questions included in the difficult category are seven, and the easy ones are seven. Questions in the straightforward category are two with a measure logit value <-0.99 . This shows that the scientific literacy instrument developed has a level of difficulty that is manageable and not too easy, so the instrument can be used for students with high abilities and students with low abilities. A good level of difficulty of questions if the questions in the test are proportional. So, the scientific literacy instrument as a whole is good. The distribution of the difficulty level of the scientific literacy test instrument can also be seen in Figure 3 below.

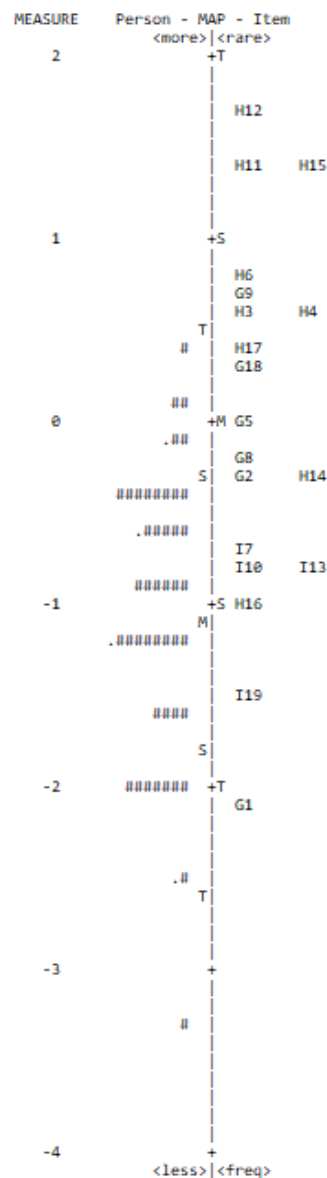


Figure 3. Map of Distribution of Difficulty Levels

4. Differentiating power

Differentiating power is the ability of questions to distinguish students who can answer questions with high ability from students who have a low ability to answer. Differentiating power is calculated by looking at the number of correct and incorrect answers from the lower and upper groups. In Rasch modeling, differentiating power analysis is carried out by analyzing individual ability levels and distinguishing the abilities of students who can answer questions and those who cannot. In addition, it can also be used to identify respondent groups based on the respondent separation index. The greater the item separation value, the better the quality of the instrument in terms of overall respondents and question items because it can identify respondent groups and item groups. Another equation to find out the grouping more precisely is the strata equation (H):

$$H = \frac{[(4 \times SEPARATION) + 1]}{3}$$

Analysis of Item Response Theory...

The separation and Person Reliability values can be seen in Figure 4. below.

SUMMARY OF 94 MEASURED Person								
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	5.9	19.0	-1.05	.58	1.00	-.04	1.09	.04
SEM	.2	.0	.08	.01	.02	.09	.06	.11
P.SD	2.3	.0	.76	.10	.22	.89	.62	1.01
S.SD	2.3	.0	.77	.10	.22	.89	.62	1.02
MAX.	11.0	19.0	.40	1.06	1.53	2.27	4.73	2.46
MIN.	1.0	19.0	-3.32	.51	.57	-2.28	.39	-1.92
REAL RMSE	.62	TRUE SD	.45	SEPARATION	.73	Person RELIABILITY	.35	
MODEL RMSE	.59	TRUE SD	.49	SEPARATION	.83	Person RELIABILITY	.41	
S.E. OF Person MEAN = .08								

Figure 4. Summary Statistic Results with Rasch Modeling

Based on Figure 4. it is known that the separation value for respondents is 0.73, so the value of $H = 1.30$ is rounded to 2, indicating that the respondent group can be divided into two groups based on the respondent separation value. Furthermore, the distribution of student abilities is also presented in Figure 5 below

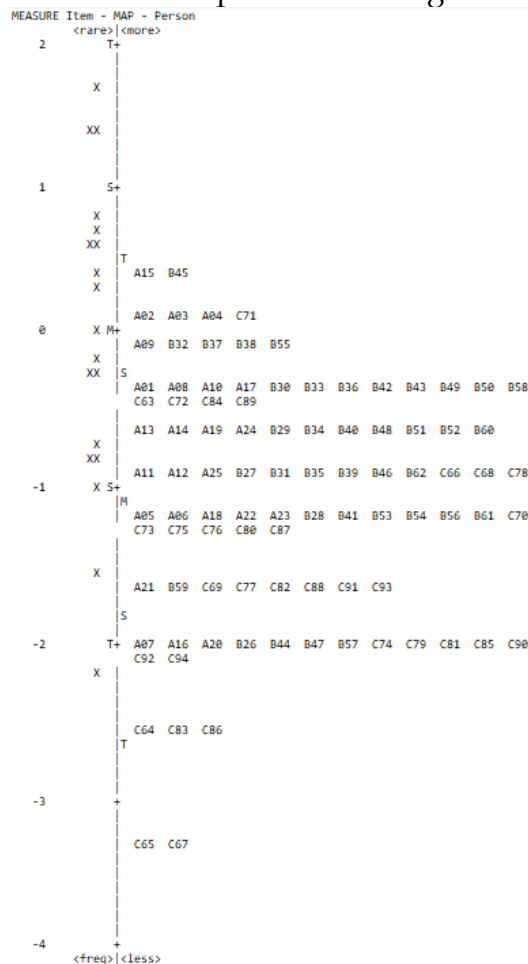


Figure 5. Map of Distribution of Student Ability Levels

The x mark on the left is the distribution of the question items, and the one on the right is the distribution of the level of students' scientific literacy abilities, from high to low. Students with the highest ability scores consist of codes A15 and B45. While students

with low abilities are students with codes C65 and C67. Based on the biased question data results, it was found that questions 3, 11, and 12 were included in the biased question category, so they had to be dropped. Therefore, based on the analysis of validity, reliability, level of difficulty, and discriminatory power, the questions that would be discarded or dropped were 3. Therefore, the final stage of the scientific literacy instrument consisted of 16 multiple-choice questions with reasons that could be used as student learning tests.

Conclusion

The developed assessment of high school physics science literacy based on Local Wisdom and Pancasila Student Profile has been proven to be of good quality after conducting empirical quality tests and analyzing them based on item response theory (IRT) using Rasch modeling. The developed science literacy assessment is declared valid because it fits the model based on the Outfit MNSQ value, Outfit ZSTD value, and Point Measure Correlation value. The assessment is declared reliable with a value of 0.92 and can group students into five groups. It has a manageable and not too easy difficulty with a measure logit value that is distributed proportionally. It has a discriminatory power with a respondent separation index of 0.73 and can differentiate students into two groups, namely groups with high ability and low ability. The final stage of the physics science literacy assessment consists of 16 multiple-choice questions with reasons that can be used as a student learning test.

This study implies that this assessment can help students relate physics concepts to local culture, increasing the relevance of learning. Assessment based on the Pancasila Student Profile supports the national education goal of producing students who are not only academically competent but also have social and cultural awareness. This study contributes to developing innovative assessment tools by integrating locality and Pancasila values, offering a new approach to physics education. Through valid and reliable assessments, this study contributes to developing students' competencies in science and literacy, which are important to face challenges in the modern era. With this approach, the study focuses on academic outcomes and the formation of students' character and cultural identity. Suggestions for further research are that it is important to provide training for teachers in implementing this assessment to understand and utilize the results to improve the teaching process. Trialing the assessment in various school and regional contexts can provide further insight into the effectiveness and adaptability of this tool.

Acknowledgement

Thanks are due to the Universitas Lambung Mangkurat which funded this research through a "Program Dosen Wajib Meneliti 2024" Grant with contract no 1090.47/UN8.2/PG/2024

References

Afnan, R., Munasir, M., Budiyanto, M., & Aulia, M. I. R. (2023). The Role of Scientific Literacy Instruments For Measuring Science Problem Solving Ability. *IJORER : International Journal of Recent Educational Research*, 4(1), 45–58. <https://doi.org/10.46245/IJORER.V4I1.271>

- Ali, L. U., & Zaini, M. (2023). Development of Interactive e-modules Based on Local Wisdom Using Android to Improve Students' Higher Order Thinking Skills (HOTS). *Jurnal Penelitian Pendidikan IPA*, 9(11), 10091–10100. <https://doi.org/10.29303/JPPIPA.V9I11.4515>
- Alifiyah, C. N., Parno, P., & Kusairi, S. (2020). Efektivitas Penggunaan UKBM Terhadap Literasi Sains Materi Alat Optik Dalam Model PjB-STEM Dengan Asesmen Formatif Pada Siswa Kelas XI MIA SMA Negeri 9 Malang. *Briliant: Jurnal Riset Dan Konseptual*, 5(4), 679. <https://doi.org/10.28926/briliant.v5i4.515>
- Alti, R. P., Lufri, L., Helendra, H., & Yogica, R. (2021). Instrumen Asesmen Berbasis Literasi Sains Tentang Materi Keanekaragaman Hayati. *Journal for Lesson and Learning Studies*. <https://ejournal.undiksha.ac.id/index.php/JLLS/article/view/34270>
- Aulia, D. M., Parno, & Kusairi, S. (2021a). Pengaruh E-module Berbasis TPACK-STEM terhadap Literasi Sains Alat Optik dengan Model PBL-STEM Disertai Asesmen Formatif. *Jurnal Riset Pendidikan Fisika*, 6(1), 7–12.
- Aulia, D. M., Parno, P., & Kusairi, S. (2021b). Pengaruh E-Modul Berbasis TPACK-STEM terhadap Literasi Sains Alat Optik dengan Model PBL-STEM Disertai Asesmen Formatif. *Jurnal Riset Pendidikan Fisika*, 6(1), 7–12. <http://journal2.um.ac.id/index.php/jrpf/article/view/16404>
- Azwar, S. (2012). Reliabilitas dan validitas. *Yogyakarta: Pustaka Pelajar*.
- Bramastia, B., Suciati, S., Nugraheni, F. S. A., Sari, M. W., Wati, I. K., Antrakusuma, B., & Masithoh, D. F. (2023). The Effectiveness of EthnoSTEM-Based Science Learning to Improve Junior High School Students' Science Literacy Ability. *Jurnal Penelitian Pendidikan IPA*, 9(SpecialIssue), 332–337. <https://doi.org/10.29303/JPPIPA.V9ISPECIALISSUE.5710>
- Chan, S. W., Ismail, Z., & Sumintono, B. (2013). A Rasch Model Analysis on Secondary Students' Statistical Reasoning Ability in Descriptive Statistics. *International Conference on Innovation, Management, and Technology Research*.
- Chasanah, N., Widodo, W., & Suprpto, N. (2022). Pengembangan Instrumen Asesmen Literasi Sains Untuk Mendeskripsikan Profil Peserta Didik. In *PENDIPA Journal of Science Education*. [ejournal.unib.ac.id. https://ejournal.unib.ac.id/index.php/pendipa/article/download/20225/9465](https://ejournal.unib.ac.id/index.php/pendipa/article/download/20225/9465)
- Dewantara, D., Mahtari, S., Misbah, M., & Haryandi, S. (2019). Student Responses in Biology Physics Courses Use Worksheets Based on Scientific Literacy. *Prisma Sains: Jurnal Pengkajian Ilmu Dan Pembelajaran Matematika Dan IPA IKIP Mataram*, 7(2), 192–197.
- Dewi, H., Negeri, S., Tumpang, K., Malang, J., & Timur, I. (2023). Penerapan Model GI-GDL untuk Meningkatkan Ketuntasan Belajar Fisika dalam Proyek Penguatan Profil Pelajar Pancasila. *Ideguru: Jurnal Karya Ilmiah Guru*, 8(2), 328–336. <https://doi.org/10.51169/IDEGURU.V8I2.519>

- Fitri. (2007). Analisis Validitas dan Reliabilitas Instrume Kinerja Akuntan Menggunakan Pendekatan Rasch Model. *Akutansi Peradaban*, III(1).
- Hidayat, R., M, A. S., & Dewantara, D. (2021). The development of electronic teaching materials on linear impulse and linear momentum to improve students' scietific literacy. *Journal of Physics: Conference Series*, 2104.
- Huda, A. A. S., & Khairina, G. P. (2023). Non-Cognitive Diagnostic Assessment of Class 7 Islamic Education Learning at Junior High School. *Tarbiyah : Jurnal Ilmiah Kependidikan*, 12(2), 127-134. <https://doi.org/10.18592/TARBIYAH.V12I2.10973>
- Huda, N., Sari, N. P., & Makaria, E. C. (2024). Perkembangan Kemampuan Bernalar Kritis Berdasarkan Teori Kognitif Agar Terwujudnya Profil Pelajar Pancasila. *Jurnal Pelayanan Bimbingan Dan Konseling*, 6(4). <https://doi.org/10.20527/JPBK.2023.6.4.11589>
- Ismail, S., Suhana, S., & Zakiah, Q. Y. (2021). Analisis Kebijakan Penguatan Pendidikan Karakter Dalam Mewujudkan Pelajar Pancasila Di Sekolah. *Jurnal Manajemen Pendidikan Dan Ilmu Sosial*, 2(1), 76-84. <https://dinastirev.org/JMPIS/article/view/388>
- Khusniati, M., Heriyanti, A. P., Aryani, N. P., Fariz, T. R., & Harjunowibowo, D. (2023). Indigenous science constructs based on Troso woven fabric local wisdom: a study in ethnosience and ethnoecology: Research Article. *Journal of Turkish Science Education*, 20(3), 549-566. <https://doi.org/10.36681/tused.2023.031>
- M Wati, M. M. (2021). The content validity of the assessment instrument on the characters of wasaka in wetland environment physics learning. *Journal of Physics: Conference Series*, 1760.
- Mansur, R. (2008). *Penilaian Hasil Belajar*. Wacana Prima.
- Matondang, Z. (2009). Validitas dan Reliabilitas Suatu Instrumen Penelitian. *Jurnal Tabularasa*, 6(1), 87-97.
- Nuraini, L., Elisa, E., Istighfarini, E. T., Prabandari, A. M., Latifa, H., Widi, L., & Hamidah, I. (2024). Diseminasi Modul Digital KOFIJATRA untuk Mendukung Penguatan Profil Pelajar Pancasila Pada Guru MGMP Fisika SMA. *Madaniya*, 5(2), 647-655. <https://doi.org/10.53696/27214834.812>
- Purwanto, P. (2016). *Evaluasi hasil belajar*. Pustaka Pelajar.
- Putra, D. S., Irawan, A., Yanto, J. P., Putra, M. M., Adli, S., Dewi, U. P., & Nasih, N. R. (2020). Relationship of Attitude with Student Learning Environment Toward Physics. *Tarbiyah : Jurnal Ilmiah Kependidikan*, 9(2), 78-86. <https://doi.org/10.18592/TARBIYAH.V9I2.3522>
- Rusmansyah, R., Leny, L., & Sofia, H. N. (2023). Improving Students' Scientific Literacy and Cognitive Learning Outcomes through Ethnosience-Based PjBL Model. *Journal of Innovation in Educational and Cultural Research*, 4(1), 1-9. <https://doi.org/10.46843/JIECR.V4I1.382>
- Santika, I. G. N. (2020). *Menggali dan Menemukan Roh Pancasila Secara Kontekstual*. Penerbit Lakeisha.

- Sari, E. N., Miriam, S., & Suyidno, S. (2023a). Developing Students' Scientific Literacy Skills in Driving Schools Through the Use of Local Wisdom-Based Physics Lesson E-module. *Berkala Ilmiah Pendidikan Fisika*, 11(1), 9–19. <https://doi.org/10.20527/bipf.v11i1.14095>
- Sari, E. N., Miriam, S., & Suyidno, S. (2023b). Developing Students' Scientific Literacy Skills in Driving Schools Through the Use of Local Wisdom-Based Physics Lesson E-module. *Berkala Ilmiah Pendidikan Fisika*, 11(1), 9. <https://doi.org/10.20527/bipf.v11i1.14095>
- Setiawan Putra, D., Irawan, A., Putra Yanto, J., Putra, M. M., Adli, S., Dewi, U. P., & Nasih, R. (2018). Uji Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif. *Tarbiyah : Jurnal Ilmiah Kependidikan*, 7(1), 78–86. <https://doi.org/10.18592/TARBIYAH.V7I1.2100>
- Setiawan Putra, D., Irawan, A., Putra Yanto, J., Putra, M. M., Adli, S., Dewi, U. P., & Nasih, R. (2021). Empirical Test of Content Validity and Construct Validity of Mathematical Self-Efficacy Scale. *Tarbiyah : Jurnal Ilmiah Kependidikan*, 10(2), 64–72. <https://doi.org/10.18592/TARBIYAH.V10I2.4341>
- Sholihah, A., Syahmani, S., & Suyidno, S. (2023). The Effectiveness of STEM Integrated Problem-Based Learning in Enhancing Student Science Literacy on Temperature and Heat Materials. *Jurnal Ilmiah Pendidikan Fisika*, 7(1), 1–9. <https://doi.org/10.20527/JIPF.V7I1.5639>
- Siswanto, J., Mahtari, S., Febriani, W., Sari, E., Lambung Mangkurat, U., Brigjen Hasan Basri, J. H., Utara, B., Banjarmasin, K., & Selatan, K. (2023). The Barriers to Developing Students' Scientific Literacy in Learning Physics of Quantities and Measurements. *Jurnal Pendidikan Sains Indonesia*, 11(2), 206–220. <https://doi.org/10.24815/JPSI.V11I2.27767>
- Suharsono, Y., & Istiqomah. (2014). Validitas Dan Reliabilitas Skala Self-Efficacy. *Jurnal Ilmiah Psikologi Terapan*, 02(01), 144–151. <http://ejournal.umm.ac.id/index.php/jipt/article/view/1776>
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch pada Assesment Pendidikan*. Trim Komunika.
- Syafitri, Y. (2023). Need Analysis for Development of Physics E-Module Integrated with Critical Reading to Enhance Critical Thinking and Scientific Literacy in Senior High School. *Jurnal Penelitian Pendidikan IPA*, 9(10), 8065–8069. <https://doi.org/10.29303/JPPIPA.V9I10.4976>
- Wati, M., Husna, A., Miriam, S., & Mahtari, S. (2018). Developing Cognitive Instruments Based on Science Literacy on Sound Materials. *1st International Conference on Creativity, Innovation and Technology*.
- Wati, M., & Mahtari, S. (2017). Pengembangan Instrumen Kognitif Fisika Siswa SMP. *Jurnal Penelitian Dan Pengkajian Ilmu Pendidikan: E-Saintika*, 1(1).
- Wati, M., Misbah, Rizki, M., Astuti, T. P., & Rizky, S. (2021). Analisis kebutuhan pengembangan instrumen karakter wasaka. *Prosiding Seminar Nasional Lingkungan Lahan Basah*, 6(April).

- Wati, M., Supriyati, Y., & Margono, G. (2015). Development of Physics Prior Knowledge Instrument for Senior High School Students on Measurement and Vector subject. *Proceeding of The 3rd International Postgraduate Conference in Science and Mathematics 2015(IPCSM'15)*.
- Wel Ningsi, A., & Mahaputra Iskandar, A. (2023). Design and Validity of Student Worksheet Integrated Scientific Literacy for The Use of Physics Practicum KIT. *Jurnal Penelitian Pendidikan IPA*, 9(1), 384-389. <https://doi.org/10.29303/JPPIPA.V9I1.2887>
- Widoyoko, E. P. (2019). *Evaluasi program pembelajaran: panduan praktis bagi pendidik dan calon pendidik*. Pustaka Belajar.
- Widoyoko, S. E. P. (2017). *Penilaian Hasil Pembelajaran di Sekolah*. Pustaka Pelajar.
- Wijayadi, A. W., Fitriyah, L. A., & Manasikana, O. A. (2023). Development of Four-Tier Instrument at First Intermediate Level Chemistry Concepts. *Tarbiyah : Jurnal Ilmiah Kependidikan*, 12(2), 149-157. <https://doi.org/10.18592/TARBIYAH.V12I2.11399>
- Yusup, F. (2018). Uji Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif. *Jurnal Tarbiyah : Jurnal Ilmiah Kependidikan*, 7(1), 17-23. <https://doi.org/10.18592/tarbiyah.v7i1.2100>