Thabiea : Journal of Natural Science Teaching
Vol. 7(1), pp. 56-78, 2024
Available online at
http://journal.iainkudus.ac.id/index.php/Thabiea
p-issn: 2580-8474, e-issn: 2655-898X

Development of science literacy competency evaluation instruments on human circulatory topic

Esty Setyo Utaminingsih[1]*, Ellianawati[2], Sri Sumartiningsih[3], FR. Wuriningsih[4], Maria Ayu Puspita[5]

[1] *Sekolah Tinggi Agama Islam Muhammdiyah Blora, Blora, Indonesia*
[2,3,5] *Universitas Negeri Semarang, Semarang, Indonesia*
[4]*Sekolah Tinggi Pastoral Kateketik Santo Fransiskus Asisi Semarang, Indonesia*
**Correspondence: estyutami@staimuhblora.ac.id*

Keywords:
Elementary school;
Evaluation;
Human blood circulatory;
Instrument test;
Science literacy;

Abstract

Teachers in elementary schools have not developed many science literacy problems so the science literacy competence of students has not been evaluated clearly and accurately. The purpose of this study is to develop a science literacy evaluation instrument and find out the feasibility of the developed instrument. The development of instruments in this study used a combination of Wilson and Oriondo and Antonio models. The instrument development steps include (1) test design, (2) test trials, and (3) test assembly. The sample number is 31 students. The sampling technique used is random sampling. The content validation test by four experts based on the Aiken assessment formula obtained a validity value of 0.9813 which means it is declared "valid" or worthy of use. Instrument validation is empirically declared "valid" and has an r value above 0.355. The value of all items is declared reliable with multiple-choice questions having a reliable value of 0.743, short-fill questions 0.744, and description questions 0.738. The results of the difference power analysis, the value of the Corrected Item-Total Correlation does not exceed the value of Cronbach's Alpha Item Deleted and has a "good" difference power interpretation value. Analysis of the difficulty level of the question items shows that the questions vary with the ideal level of difficulty with an easy comparison: medium: difficult which is 3: 5: 2. It can be concluded that overall the science literacy question items are in good condition and feasible for measuring students' science literacy competence.

To cite this article:
Utaminingsih, E.S., Ellianawati, Sumartiningsih, Sri., Wuriningsih, FR., Puspita, M.A. (2024). Development of science literacy competency evaluation instruments on human circulatory topic. *Thabiea : Journal of Natural Science Teaching*, 7(1), 56-78.

Introduction

Education aims to create superior human resources to face 21st-century education in the era of society 5.0 that can be globally competitive (Adnan et al., 2021). 21st-century education leads students to have skills in solving problems (Gurses et al., 2015). Students must be able to solve complex problems with digital media support or utilize technology (Utaminingsih, Raharjo, & Ellianawati, 2023). In addition, they are also expected to use scientific knowledge and perform scientific reasoning in decision-making in everyday life (Kalkan et al., 2020). Scientific literacy is the ability to use scientific knowledge, identify problems (Flores, 2018), and draw conclusions based on scientific evidence in making decisions to solve problems through daily activities (Lawless et al., 2018).

Scientific literacy aided students in comprehending fundamental scientific principles applicable to everyday life (Bucchi & Saracino, 2016). It was particularly crucial in teaching complex concepts such as the human circulatory system, which is not only a cornerstone of biological education but also essential for a robust understanding of health (Trémolière & Djeriouat, 2021). The introduction of the circulatory system concept at this level helped students develop a strong understanding from an early age, which facilitated the comprehension of more advanced scientific concepts at higher educational levels (Utaminingsih, Ellianawati, Sumartiningsih, et al., 2023). A longitudinal study by Afriana et al. (2016) found that elementary school students taught complex biological concepts, including the circulatory system, showed more excellent proficiency in scientific literacy tests during middle school compared to their peers who received less intensive science education in their early years. A science education can lead to sustained academic benefits.

Science literacy is a competency that must be mastered in the 21st century (Li et al., 2020). Science literacy is used as a guideline to answer challenges in the global era (Ahied et al., 2020), so it is expected that students can face the demands of the times as Problem solvers (Febriyanti & Sari, 2022) as well as individuals who are creative, competitive, innovative, and have character according to technology and science (Utami &; Desstya, 2021). Science literacy learning in educational institutions aims to train students to have the ability to apply the knowledge competencies obtained while in school (Amaringga et al., 2021), then make provisions for living well in the community by making the right decisions to solve problems (Qadar et al., 2022) through the provision of knowledge, they have (Ding, 2022).

Referring to the importance of science literacy competencies that students must master, measuring how far they have competence is necessary (Sultan et al., 2021). However, evaluation data by The Programme for International Student Assessment (PISA) from 2000 to 2018 shows that the science literacy of students in Indonesia still needs to be improved (McComas, 2019; Schleicher, 2018). The PISA assessments, held every three years, evaluate 15-year-olds' reading, mathematics, and science literacy abilities, focusing on their capacity to apply knowledge to real-life situations. In 2018, Indonesian students scored an average of 371 in reading, which is well below the OECD average of 487. Similarly, the average scores in mathematics and science were also below the international average, indicating a pervasive challenge across science literacy. A survey by Trend in International Mathematics and Science Study (TIMSS) in 2015 shows that the science literacy ability of students in Indonesia is in the deficient category (Suparya et al., 2022). It is undoubtedly a problem that needs attention (Li et al., 2020), considering that science literacy competence is one of the primary needs of learners in the 21st century (Utaminingsih, Ellianawati, et al., 2023; Kasse *et al.*, 2022). Some of the results of measuring science literacy are in line with the results of interviews at Supriyadi Semarang Elementary School, in which teachers do not have the opportunity and competence to develop science literacy problems (Muthmainnah & Istiyono, 2019).

During the Supriyadi Semarang Elementary School interview, the teacher said the evaluation results were carried out some time ago and no evaluations have been made using systematic instruments to obtain accurate results. Referring to the interview results and considering the importance of science literacy competence, measuring students' initial science literacy competence is necessary to determine further actions to improve it (Bahri et al., 2021). Evaluation can be done through a science literacy competency test instrument by the indicators to be evaluated (Rokhmah et al., 2017). Therefore, it is necessary to develop instruments to evaluate students (Utaminingsih & Puspita, 2023), particularly to evaluate science literacy competence.

The evaluation of science literacy competency instruments cannot stand alone but can be combined with subjects (Ulva et al., 2021). Science learning can prepare aspects of science literacy in students so that they will have science literacy skills (Ahied et al., 2020). In addition, it provides space to increase insight, develop skills, and utilize technological sophistication in everyday life (Fakhriyah et al., 2017). The results of interviews with teachers at Supriyadi Semarang Elementary School also found that one of the materials that is difficult for students to master is science learning on human blood circulation. Therefore, it is the right step for developing science literacy competency evaluation instruments to collaborate with science learning, specifically on human blood circulation, following the school's problems.

Evaluation in learning is essential to evaluate the achievement of the competencies (Pomalato et al., 2021) and the extent to which students can understand the material taught (Noroozi & Mulder, 2017). Evaluation is carried out to collect and process information to assess the achievement of student learning processes and outcomes, guided by instruments (Bartimote-Aufflick et al., 2016). The instruments developed are also based on the indicators to be achieved (Chung et al., 2022). The importance of evaluation is not just products or results but a series of activities demonstrating the quality assessed (Septiani et al., 2022).

The research on the development of a scientific literacy instrument by Rusilowati in 2018 focused on different indicators from those typically measured in other studies. The investigated dimensions in Rusilowati's research included the knowledge of science, the investigative nature of science, science as a way of thinking, and the interaction between science, environment, technology, and society. In contrast, the current study aimed to develop an instrument that would assess new indicators. These included: (1) identifying valid scientific opinions; (2) conducting effective literature searches; (3) creating accurate charts based on relevant data; (4) solving problems using quantitative skills, including basic statistics; (5) understanding and interpreting the results of statistical analysis; (6) drawing conclusions and making predictions based on quantitative data; and (7) evaluating scientific information. These indicators signify a shift towards assessing practical and applied scientific skills, reflecting a more comprehensive approach to scientific literacy. Based on the results of interviews and the problems found, this study aims to (1) develop instruments for measuring scientific literacy competence in human blood circulation and (2) Analyze the feasibility level of science literacy instruments.

Method

This research is a *Research and Development* (R&D). The research design used in this study was a combination of the Wilson and Oriondo and Antonio models (Astuti et al., 2020). The steps of instrument development in the form of tests using modifications, include: (1) test design, (2) test trials, and (3) test assembly (Jiananda, 2017). This development model makes

it easy for the implementation of research to evaluate the achievement of the products produced. The design in the study is presented in Figure 1.
study is presented in Figure 1.

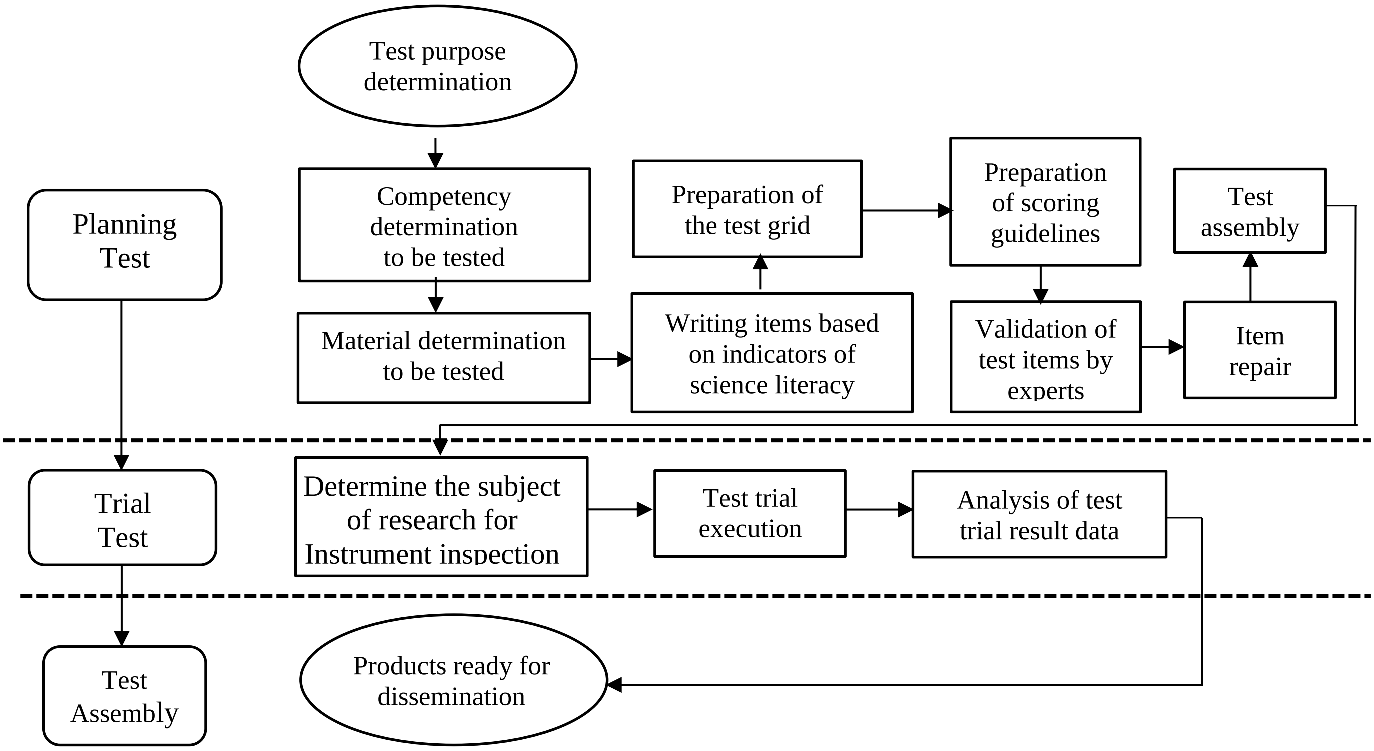


Figure 1. Research Design

The indicators of the three scientific literacy competencies that will be used as a reference for problem development consist of seven indicators. The indicators of science literacy according to PISA (Adnan *et al*., 2021) are presented in Table 1.

Table 1. Competence and indicators of science literacy

Scientific Competence Science Literacy	Indicator
1. Identify scientific issues (problems)	a. Identify valid scientific opinions (e.g., opinions/theories to support hypotheses).
	b. Conduct effective literature searches (e.g., evaluating the validity of sources and distinguishing source types).
	c. Understand the elements of research design.
	d. Accurate data analysis.
2. Explaining scientific phenomena	a. Develop accurate charts based on relevant data.
	b. Solve problems using quantitative skills, including basic statistics (e.g., calculating averages, probabilities, percentages, frequencies).
	c. Understand and interpret the results of statistical analysis.
3. Using scientific evidence	a. Conclude and make predictions based on quantitative data.
	b. Evaluate scientific information.

The results of developing questions based on scientific literacy indicators led to human circulatory material that experts then validated. The validators were experts in their respective fields relevant to the research topic. In this case, the validators consisted of two university lecturers, one specializing in Natural Science Education and the other in Primary School Teacher Education, along with two elementary school teachers from fifth grade. Test the validity of the content by experts using the Aiken formula. Aiken formulated Aiken's V formula for calculating the content-validity coefficient based on an expert panel's assessment of an item in terms of the degree to which the item represents the evaluated construct. The assessment is carried out by giving a checklist from the category "Irrelevant" with a score of "1" to "Very Relevant" with a score of "5". Aiken's V formula (Aiken, 1985) is as follows.

$$\frac{\Sigma S}{[n(c-1)]} \qquad (1)$$

Information:
S = r – lo
r = number given by the appraiser
lo = lowest validity assessment number
n = number of appraisers
c = highest validity assessment number

Determining the validity of each item assessed can be analyzed by looking at the value of the validity coefficient V contained in the Validity Coefficient Value Table presented in Table 2.

Table 2. Value of Validity Coefficient*

Raters	**Number of Ranting Categories**							
	2		**3**		**4**		**5**	
	V	**P**	**V**	**P**	**V**	**P**	**V**	**P**
2							1	0.040
3							1	0.008
3			1.00	0.037	1	0.016	0.92	0.032
4					1	0.004	0.92	0.032
4			1.00	0.012	0.92	0.020	0.88	0.024
5			1.00	0.004	0.93	0.006	0.90	0.007
5	1.00	0.031	0.90	0.025	0.87	0.021	0.80	0.040

*****Aiken, 1985

The experts who validated the question items were four with five scoring categories. Therefore, this study's content validity analysis results are guided by the fourth raters column with $p < 0.05$ in the second row and the *number of ranting categories* (scale) in the fifth column of 0.88. In addition, the level of validity analyzed by Aiken's equation should refer to the Kappa Statistic validity classification category.

Kappa statistic or *interrater reliability* is an evaluation used to test agreement between two people (raters/observers) on categorical variables. Several researchers, such as Wynd et al. (2003), use this technique and CVI and kappa multi-rater to validate their developed scale content. They argue that statistical kappa is a vital supplement, not a CVI substitute. Kappa provides information about the extent of agreement beyond possibility (Polit & Beck, 2006). If there are more than two raters, the Kappa multi-rater technique can be used. Kappa Cohen's statistical evaluation of interrater reliability generally ranges from 0 to 1.0, where a large number means better reliability and a value close to or less than zero indicates that the deal was caused by chance alone (Polit et al., 2007). Landis & Koch (1977) provide assessment guidelines on the Kappa statistic presented in Table 3.

Table 3. Kappa Statistical Assessment Guidelines*

Kappa	*Interpretation*
< 0	*Poor Agreement*
0.0 – 0.20	*Slight Agreement*
0.21 – 0.40	*Fair Agreement*
0.41 – 0.60	*Moderate Agreement*
0.61 - 0.80	*Substantial Agreement*
0.81 – 1.00	*Almost perfect Agreement*

*Landis & Koch, 1977

The development of question items after expert validation and questions declared feasible according to the Aiken formula and Kappa Statistics guidelines would be tested on fifth-grade students, at Supriyadi Semarang Elementary School. The sampling technique in determining the sample used in this study is by random sampling technique by Moleong (2017). The instruments were tested on 31 students and then analyzed using the SPSS 25 Statistics Program. The data will be analyzed for the question items' validity, reliability, differentiation, and level of difficulty.

The validity test is carried out on the question items that have been developed. The next validity test is empirical validity. The practical validity test of the question items was carried out using the SPSS 25 Statistic Program. Validity testing correlates each indicator item's score with the total score. The significance level used is 0.05. The test criteria are: H_0 is accepted if r counts > r table, which means (the measuring instrument used is valid or valid). H_0 is rejected if the statistical r ≤ r table means that the measuring device used is invalid or invalid). The next step is to interpret the score of each item with the r table. The interpretation of r table values is presented in Table 4.

Table 4. Distribution of Validity Values r 5% Significance*

N	The Level of Significance (5%)
21	0.433
22	0.432
23	0.413
24	0.404
25	0.396
26	0.388
27	0.381
28	0.374
29	0.367
30	0.361
31	0.355
32	0.349

*Janna & Herianto, 2021

Reliability indicates the reproducibility of item evaluations if the item is given to other samples taken from the same population or the reproducibility of evaluations on other occasions (Bashooir & Supahar, 2018). Reliability shows how far an evaluation is carried out repeatedly but produces the same information and is related to the consistency of assessment results (Moleong, 2017). Reliability analysis was performed with the SPSS 25 Statistic Program. The reliability obtained from the results of data analysis is reviewed from the value of Alpha Cronbach. Test instruments are reliable when the Alpha Cronbach coefficient value ≥ 0.70 or at least meets the interpretation criteria of fixed/good reliability (Payadnya & Jayantika, 2018). Furthermore, reliability levels are classified by interpretation criteria in Table 5.

Table 5. Reliability Interpretation*

Correlation Coefficient	Correlation	Reliability Interpretation
0.90 ≤ r ≤ 1.00	Very high	Very fixed/ very good
0.70 ≤ r ≤ 0.90	High	Fixed / good
0.40 ≤ r ≤ 0.70	Keep	Quite fixed/ good enough
0.20 ≤ r ≤ 0.40	Low	Not fixed/ bad
r < 0,20	Very low	Not fixed/ bad

*Rofiyadi & Handayani, 2021

The next stage is a different power test. The difference power test is carried out to determine the extent to which the questions can distinguish students with high scores and low scores. This test aims to determine the ability of question items to find different powers, namely to determine whether question items that have been developed can distinguish high-achieving groups (upper group) from low-achieving groups (lower group) among test participants. The difference considered still sufficient for a question is when it is equal to or greater than 0.30 or the value of Cronbach's Alpha coefficient if Item Deleted does not exceed the value of Cronbach's Alpha coefficient (Payadnya & Jayantika, 2018). In addition, the differentiating power interpretations are classified in Table 6.

Table 6. Interpretation of Differentiating Power*

Interval	Interpretation
0.00 – 0.19	Poor
0.20 – 0.39	Satisfactory
0.40 – 0.69	Good
0.70 – 1.00	Excellent

*****Nani, 2021

The development of question items also goes through the difficulty level test stage. The difficulty level of the question item shows the possible value of the number of respondents who can answer the question item correctly. The ideal question has a medium difficulty level (Susanto et al., 2015). The difficulty level of the question items can be calculated using the following formula.

$$IK = \frac{\overline{X}}{SMI} \qquad (2)$$

Information:
IK : the difficulty index of the question item
$\overline{X}$: the average score of each question item
SMI : ideal maximum score (maximum score)
The difficulty index of a question item is interpreted in categories, as shown in Table 7.

Table 7. Interpretation of Difficulty Levels

Value	Interpretation of the Difficulty Index
$IK = 0.00$	Too difficult
$0.01 \le IK \le 0.30$	Difficult
$0.31 \le IK \le 0.70$	Keep
$0.71 \le IK \le 0.99$	Easy
$IK = 1.00$	Too easy

The ideal question criteria are between 0.31 and 0.70, with a medium interpretation category. This refers to Nani's (2021) statement (Susanto et al., 2015). The results of the difficulty level analysis in Table 3.8 can be used as a reference for choosing the ideal question. However, the selection needs to consider the results of validity, reliability, and differentiation.

Instruments declared valid, reliable, have different strengths and a good difficulty level are then included in digital modules as e-modules. This instrument can be used both on paper and electronically. Electronic test instruments utilize Microsoft PowerPoint with additional features, namely iSpring Suite 11. Instruments incorporated into digital modules are called quizzes.

Results and Discussion

Results of Question Point Development

The development of scientific literacy problems leads to evenly distributed human circulatory material based on three competencies and seven indicators. The number of questions developed consists of twenty questions. The distribution of question items is based on competence and indicators of scientific literacy and the distribution of topics, namely the heart, blood vessels, and lungs. Variative questions include multiple choice (MC), short fills, and descriptions. The complete distribution of question items is presented in Table 8.

Table 8. Distribution of Question Items Based on Indicators and Topic

Ability Aspect	Competency Achievement Indicators	Question Indicator	Question Number	Question Type
Identify scientific issues (problems)	Identify valid scientific opinions	Present infographics about cases of smoking-related diseases,	1	MC
		then ask learners to identify the	2	MC
		cause of heart disease	3	MC
	Carry out an effective literature search	Search for information about the heart and blood vessels	4	MC
			5	MC

Ability Aspect	Competency Achievement Indicators	Question Indicator	Question Number	Question Type
Explaining scientific phenomena	Develop accurate charts based on relevant data	Complete the circulatory chart	20	Essay
	Solve problems using quantitative skills, including basic statistics (e.g., calculating averages, probabilities, percentages, and frequencies).	Identify a healthy heart rate.	11 12	Short Fill Short Fill
		Identify the comparison of the size of the heart compared to the rest of the body	13 14	Short Fill Short Fill
		Mentions the length of the entire blood vessel	6	MC
	Understand and interpret the results of statistical analysis.	Analyze healthy lung function and capacity based on the data presented	16 7	Essay MC
		Analyze heart function based on infographics	15	Short Fill
Using scientific evidence Identify scientific issues (problems)	Conclude and make predictions based on quantitative data	Analyze the average heart rate of people who exercise.	8	MC
		Making a conclusion based on information about the presented parts of the blood vessels.	9	MC
		Predicting the lung condition of smokers in industrial areas and rural areas.	17	Essay
		Predicting smokers' heart conditions.	18	Essay
	Evaluate scientific information	Formulate types of blood vessels and their functions	10	MC
		Complete the parts of the heart and their functions.	19	Essay

Instrument Quality Based on Content Validity

The design of the science literacy assessment instrument approved by the supervisor is then validated with expert validation by four validators, consisting of two lecturers and two fifth-grade elementary school teachers. The results of the expert validity analysis of science literacy questions are presented in Table 9.

Table 9. Results of Validity of Science Literacy Question Items

Question Number	V Value	CVI Aiken	Category KS Validity Status
1	1	Valid	Almost Perfect
2	0.9375	Valid	Almost Perfect
3	0.9375	Valid	Almost Perfect
4	1	Valid	Almost Perfect
5	0.9375	Valid	Almost Perfect
6	1	Valid	Almost Perfect
7	1	Valid	Almost Perfect
8	1	Valid	Almost Perfect
9	1	Valid	Almost Perfect

Question Number	V Value	CVI Aiken	Category KS Validity Status
10	0.9375	Valid	Almost Perfect
11	1	Valid	Almost Perfect
12	1	Valid	Almost Perfect
13	1	Valid	Almost Perfect
14	1	Valid	Almost Perfect
15	1	Valid	Almost Perfect
16	1	Valid	Almost Perfect
17	1	Valid	Almost Perfect
18	1	Valid	Almost Perfect
19	1	Valid	Almost Perfect
20	0.875	Invalid	Almost Perfect
Final V Value	**0.9813**	**Valid**	**Almost Perfect**

Based on Table 9, the results of expert validation of the science literacy assessment instrument obtained a final V value of 0.9813. The results are then reviewed from the value of the Aiken validity coefficient with the number of raters, as many as 4 (four) experts, and the questionnaire scale used; there are 5 (five) scales, with $p < 0.05$. Based on the validity level analysis results with Aiken's equation, the V value of 0.9813 is included in the "valid" category. However, when reviewed on a per-item basis, 1 (one) Item is "invalid" according to the CVI table. The Item explains scientific phenomena in the question item "developing an accurate chart based on relevant data ."However, Kappa Statistic is another reference to categorize a product's validity (Polit et al., 2007). According to the assessment guidelines in Kappa Statistics (Landis & Koch, 1977), the V value of 0.875 falls into the "Almost Perfect" category. Therefore, question point number twenty, with a V value of 0.875 according to the Kappa Statistics assessment guidelines, is still allowed.

Referring to the Kappa Statistics assessment guidelines, based on expert validation, the science literacy assessment instrument has met the declared "valid" with a validity value of 0.9813. This valid result aligns with research on developing science literacy problems by Kusnanto et al. (2021). The difference in research on the development of this instrument is that the questions developed in the study were only 15, while 20 questions were asked. In addition, the development of questions in addition to holding on to the material is also based on seven indicators of science literacy competence according to PISA, while the development of science literacy instruments in the research of Kusnanto et al. (2021) covers the cognitive domains C1 (knowledge) to C5 (evaluation), namely knowledge (C1), understanding (C2), application (C3), analysis (C4) and evaluation (C5).

Instrument Quality Based on Empirical Validity

After being declared "valid" by experts, empirical validity tests were carried out by testing question items on students of Supriyadi Semarang Elementary School students. This trial was carried out on 31 fifth-grade students. The results of the students' work were then analyzed using the SPSS 25 Statistics program. An example of the results of working on science literacy questions by students is presented in Figure 2.

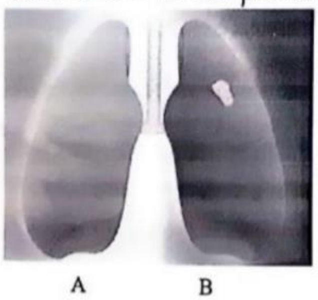
IV. Jawablah pertanyaan di bawah ini dengan lengkap dan jelas!

16. Dikatakan paru-paru masih berfungsi dengan baik apabila dalam 6 (enam) menit jarak yang dapat ditempuh oleh seseorang sesuai dengan usianya tersaji pada tabel berikut:

Usia	Jarak
12-20 tahun	300-500 meter
21-44 tahun	400-600 meter
45-54 tahun	300-500 meter
Di atas 54 tahun	200-300 meter

0

Saat ini usia Pak Burhan adalah 45 tahun. Dua tahun yang lalu, selama 6 menit Pak Burhan dapat berjalan 350 meter. Apakah kondisi paru-paru Pak Burhan saat itu berfungsi dengan baik? Mengapa?
baik, krn tidak melebihi jarak 500 m

17. 

3

Perokok memiliki kemungkinan 12-13 kali lebih tinggi untuk meninggal akibat penyakit paru-paru. Satu batang rokok menyebabkan umur seseorang memendek 12 menit. Manakah diantara paru-paru A dan B yang merupakan paru-paru perokok? Mengapa? Mengapa mereka memiliki kemungkinan meninggal lebih tinggi?
B, krn di paru² tsb sedikit hitam dan mempunyai lubang
Krn mereka kemungkinan 12-13 x lebih tinggi untk meninggal

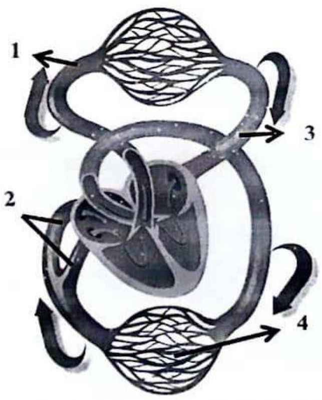
18. Jika seseorang merokok 1 batang/hari: memiliki risiko penyakit jantung 1.48x. Jika 5 batang/ hari: memiliki risiko 1.58x. Jika 20 batang/hari: memiliki risiko 2.04x.
   a. Jika seseorang merokok 1 batang setiap 5 jam, maka berapa kira-kira risiko yang dia miliki selama 24 jam? 1.58 x
   b. Jika seseorang merokok 1 bungkus di tambah 8 batang sehari Berapa kali risiko yang dimiliki oleh seorang perokok? (keterangan: 1 bungkus rokok berisi 12 batang) 2.04 x

19. Sebutkan dan jelaskan ruang jantung yang ditunjukkan oleh nomor satu sampai dengan nomor empat (boleh memilih dua nomor)!

4

1. Serambi kanan : berfungsi sebagai menerima darah yg mengandung CO_2
2. " Kiri : berfungsi sebagai menerima darah yg mengandung O_2

20. Sebut dan jelaskan bagian-bagian dari sistem peredaran darah di bawah ini!



0

Figure 2a. Question Item Test Results Were Tested

IV. Answer the questions below entirely and clearly!

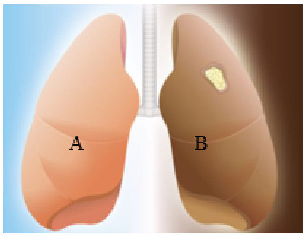
16. The lungs still function correctly if within 6 (6) minutes. The distance a person can cover according to their age is presented in the following table:

Age	Distance
12-20 years	300-500 meters
21-44 years	400-600 meters
45-54 years	300-500 meters
Up to 54 years	200-300 meters

Currently, Mr. Burhan is 45 years old. Two years ago, in 6 minutes, Mr. Burhan could walk 350 meters. Were Mr Burhan's lungs functioning well at that time? Why?

Answer: Good, because it does not exceed a distance of 500 meters.

17.

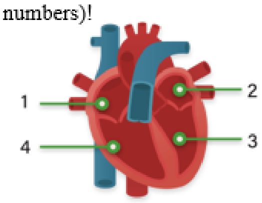


Smokers are 12-13 times more likely to die from lung disease. One cigarette causes a person's lifespan to shorten by 12 minutes. Which of lungs A and B are smoker's lungs? Why? Why do they have a higher chance of dying?

Answer:

B, because the lungs are a little black and have holes. Because they are 12-13x more likely to die.

18. If someone smokes one cigarette/day, the risk of heart disease is 1.48x. If five cigarettes/day: has a risk of 1.58x. If 20 cigarettes/day: has a risk of 2.04x.
   a. If a person smokes one cigarette every 5 hours, then approximately what risk does he or she have over 24 hours?
      Answer: 1,58
   b. If someone smokes one pack plus eight cigarettes a day, how many times the risk does a smoker have? (remarks: 1 pack of cigarettes contains 12 cigarettes)
      Answer: 2.04
19. Name and explain the heart chambers indicated by numbers one to four (you can choose two numbers)!



Answer:

1. The right atrium functions to receive blood containing CO_2
2. The left atrium functions to receive blood containing O_2

20. Name and explain the parts of the circulatory system below!

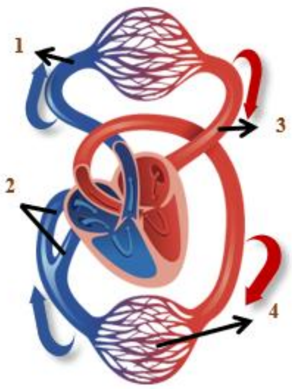


Figure 2b. Question Item Test Results Were Tested in English

Evaluation of instrument quality, in addition to being obtained from the validity of the content, empirical validity tests are also carried out by testing the validity of the question items both on science literacy questions. The validity value of the question items was analyzed using the SPSS 25 Statistics program. The results of the validity test of science literacy questions are presented in Table 10.

Table 10. Results of Validity of Science Literacy Question Items

Question Number	Value	Criterion
1	0.558	Valid
2	0.500	Valid
3	0.518	Valid
4	0.687	Valid
5	0.756	Valid
6	0.563	Valid
7	0.610	Valid
8	0.417	Valid
9	0.698	Valid
10	0.487	Valid
11	0.489	Valid
12	0.372	Valid
13	0.722	Valid
14	0.806	Valid
15	0.665	Valid
16	0.694	Valid
17	0.750	Valid
18	0.722	Valid
19	0.524	Valid
20	0.399	Valid

Table 10 shows that of the twenty question items developed to assess science literacy competence, from multiple-choice questions, short fills, and descriptions tested to thirty-one students, each question item is classified as "valid." Analysis of validity criteria, in addition to being shown from the results of the correlation analysis of the significance of SPSS Statistics, and also based on the distribution of table r values with a significance of 5% adjusted to the number of respondents. Based on the r table with thirty-one respondents, each question item is declared valid if each question item has an r value above 0.355 (Payadnya & Jayantika, 2018). All items of science literacy questions have an r value above 0.355, so they are declared "valid." This means that the question items tested are in good condition for evaluation (Herlina et al., 2020). These question items can generally be evaluated (Pratiwiningtyas et al., 2017).

Instrument Reliability

Reliability tests are conducted on science literacy instruments with critical and independent reasoning characteristics. The analysis was carried out using the SPSS 25 Statistic Program. Science literacy test instruments, critical reasoning characters, and independent students are said to be reliable if the value of the correlation coefficient ≥ 0.70 or at least meets the interpretation criteria of fixed / good reliability (Payadnya & Jayantika, 2018). The trial was conducted on thirty-one Supriyadi Semarang Elementary School grade V students. The results of the reliability test are presented in Table 11.

Table 11. Results of Science Literacy Instrument Reliability Test

Question Type	Question Number	Value *Alpha Cronbach*	Criterion
Multiple choice	1-10	0.743	Fixed / good
Short Fill	11-15	0.744	Fixed / good
Essay	16-20	0.738	Fixed / good

The analysis showed that the instrument developed could be categorized as "reliable" because the Alpha Cronbach score obtained was included in the "fixed/good" category tested on thirty-one students. Cronbach's Alpha indicates the interaction between respondents and items (Arsi, 2021). It means that the interaction between respondents and items in both tests is included in the excellent category (Laliyo et al., 2020), so that the test instrument can be tested on respondents.

Differentiating Power

The differentiating power test was carried out on the instrument to assess the competence of science literacy and the character of Pancasila students. The purpose of this test is to determine the level of ability of question items in finding different powers, namely to determine whether the question items that have been developed can distinguish high-achieving groups (upper group) from low-achieving groups (lower group) among test participants. The difference considered to still exist for a question is when it is equal to or greater than 0.30 (Payadnya & Jayantika, 2018). The results of the difference power test analysis of the developed science literacy question items are presented in Table 12.

Table 12. Results of the Differentiating Power Analysis of Science Literacy Questions

Question Number	Corrected Item-Total Correlation	Cronbach's Alpha Item Deleted	Interpretation
1	0.490	0.724	Powerful/satisfying
2	0.437	0.730	Powerful/satisfying
3	0.443	0.726	Powerful/satisfying
4	0.634	0.713	Powerful/satisfying
5	0.714	0.710	Powerful/satisfying
6	0.505	0.726	Powerful/satisfying
7	0.559	0.724	Powerful/satisfying
8	0.362	0.737	Powerful/satisfying
9	0.644	0.711	Powerful/satisfying
10	0.409	0.729	Powerful/satisfying
11	0.333	0.742	Powerful/satisfying
12	0.291	0.756	Weak/unsatisfactory
13	0.616	0.687	Powerful/satisfying
14	0.734	0.669	Powerful/satisfying
15	0.552	0.703	Powerful/satisfying
16	0.534	0.685	Powerful/satisfying
17	0.688	0.693	Powerful/satisfying
18	0.595	0.675	Powerful/satisfying
19	0.434	0.715	Powerful/satisfying
20	0.331	0.734	Powerful/satisfying

Table 12 shows that items 1 to 10 science literacy questions have substantial or satisfactory interpretations. This is because the table shows the literacy question item of the Corrected Item-Total Correlation value > 0.3. In addition, the assessment is also reviewed based on the value of Cronbach's Alpha Item Deleted, which must be less than the value of Cronbach's Alpha. Question items 1 through 10 qualify as having a "strong" or "satisfactory" interpretation because Cronbach's Alpha Item Deleted scores < 0.743 (Cronbach's Alpha scores of science literacy items 1-10).

Short Fill question items number 11 to 15, there is one item item that has a weak or unsatisfactory interpretation with a Corrected Item-Total Correlation value of < 0.3 and a Cronbach's Alpha Item Deleted value of more than 0.744 (Cronbach's Alpha value of science literacy items number 11-15), namely in question point number 12. However, question point number 12 is not omitted from the instrument, and this is because it is used as an evaluation of the science literacy aspect of "explaining scientific phenomena" on the indicator "solving problems using quantitative skills, including basic statistics." On the other hand, Nani (2021) states that the interval 0.20-0.39 has a different power interpretation value of "sufficient". This means that the Item can still be used. In addition, the validation results by validators, science literacy question number 12 have a V value of "1" or "valid," so this is considered, and question item number 12 is not eliminated from the science literacy instrument. As for question items number 11, 13, 14, and 15 already have an interpretation value of "strong" or "satisfactory".

The description question items from 16 to 20 have a "strong" or "satisfactory" interpretation value. This is because when viewed from the value of the Corrected Item-Total Correlation, all items have a value of more than 0.3. The value of Cronbach's Alpha Item Deleted items 11 through 14 is not more than 0.738, but point 15 exceeds 0.738. It can be concluded that question items 16 to 20 have an interpretation value of "strong" or "satisfactory" difference power. It has a substantial difference in power value. This indicates that the developed question items can be used to evaluate learning achievement (Bahri et al., 2021).

The Difficulty Level of The Question Items

The difficulty level of the question item shows the possible value of the number of respondents who can answer the question item correctly. Difficulty level analysis determines if the problem is classified as easy or difficult (Solichin, 2017). Good questions could be more challenging or challenging (Payadnya & Jayantika, 2018). The results of the difficulty level analysis of science literacy questions are presented in Table 13.

Table 13. Results of the Difficulty Level Analysis of Science Literacy Instruments

Question Type	Question Criteria	Percentage (%)	Question Number
Multiple Choice	Easy	30	2 and 5
	Keep	50	1, 3, 4, 9, and 10
	Difficult	20	6 and 8
Short Fill	Easy	0	-
	Keep	80	12
	Difficult	20	11,13,14, and 15
Essay	Easy	0	-
	Keep	40	17 and 18
	Difficult	60	16, 19, and 20

The results of the difficulty level analysis in Table 13 are distributed from questions classified as "easy," "medium," and "difficult." There is a significant difference in the distribution of questions on each criterion. In the science literacy question point, no questions are "too easy" or "too difficult. "Based on these results, the instrument for measuring science literacy competence is good. This is supported by the statement by Solichin (2017) that a good question is a question that is "not too easy" or "not too difficult." More easy problems are needed to stimulate students to increase efforts to solve them. Conversely, questions that are too difficult will cause students to become discouraged and not have the enthusiasm to try again because it is beyond their reach.

According to Susanto et al. (2015), to obtain good quality questions, in addition to meeting validity and reliability tests, question items must also balance the difficulty of the question items. Table 4.20 shows that science literacy multiple-choice questions have ideal proportions, namely with a percentage of 30%:50%:20% or 3:5:2 with easy, medium, and complex criteria. It is supported by the statement of Novalia & Syazali (2014) that the ideal problem has a proportional distribution, namely with an easy-medium-difficult ratio of 3-5-2. Short questions and essays only meet two criteria for good questions, which are only medium and challenging, so the questions are close to ideal. In line with the findings of Rohana et al. (2018), In the development of tests to measure scientific literacy competence, the test questions developed have moderate and difficult criteria, which are said to be close to ideal. The question items are still used to measure the science literacy of students because, according to Rusilowati (2018), the quality of the questions, in addition to being measured by the level of difficulty, while the difference in the question items is relatively good. Magdalene *et al*. (2021) also said that the quality of the questions is not only seen from the question items' difficulty level but also from the test values of validity, reliability, and differentiation.

The analysis results on the instrument can be used as a reference for choosing the ideal problem. However, the selection needs to consider the results of validity, reliability, and differentiation. Based on the validity test consisting of validity tests, reliability tests, difference power tests, and difficulty tests, all questions are declared "valid" and "reliable," have strong differentiation, and the difficulty level of the question items have a good distribution. Therefore, there is no need for the question items to be eliminated. The instrument can measure scientific literacy competence on large-scale tests based on the results above.

Test Instruments Submitted to Digital Modules

The instruments developed have been declared valid and reliable, have different strengths, and a good level of difficulty. The next step is combined into a digital module (e-module), a STEAM-based e-module in the form of .apk application file developed by Utaminingsih (2023). This is adjusted to technological developments where students use many devices in daily and school activities (Zulfahrin et al., 2019). In addition, combining test instruments (from now on referred to as quizzes) with modules creates innovations so that students will feel different experiences and become more interested when doing quizzes. This is supported by (Ilmi et al., 2021); in their findings, students will find learning more enjoyable, and the measurement of learning outcomes will feel like playing when tests are included in digital teaching materials. The results of the conversion of science literacy competency questions combined into digital modules are presented in Figure 3.

Figure 3. Quiz Display in E-Module

Figure 3 displays a science literacy measurement instrument converted into a different format and incorporated into an Android-based e-module. The science literacy competence instrument in the e-module is named Sequizi, to attract any student's attention. Students can access the quiz after learning the topic presented. The quiz wa divided into Quiz A, B, and C. The complete display of the quiz developed and combined with the e-module is presented in Figure 4.

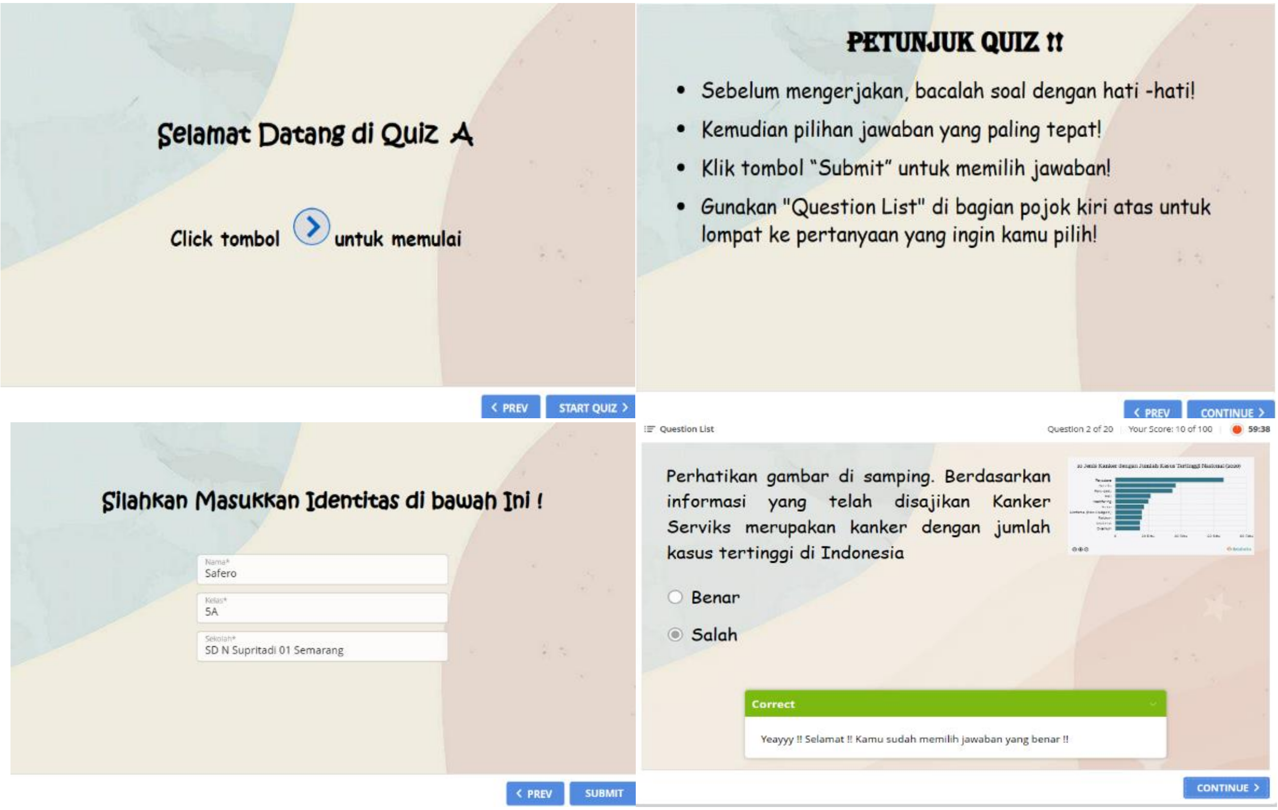


Figure 4a. Quiz Display When Running

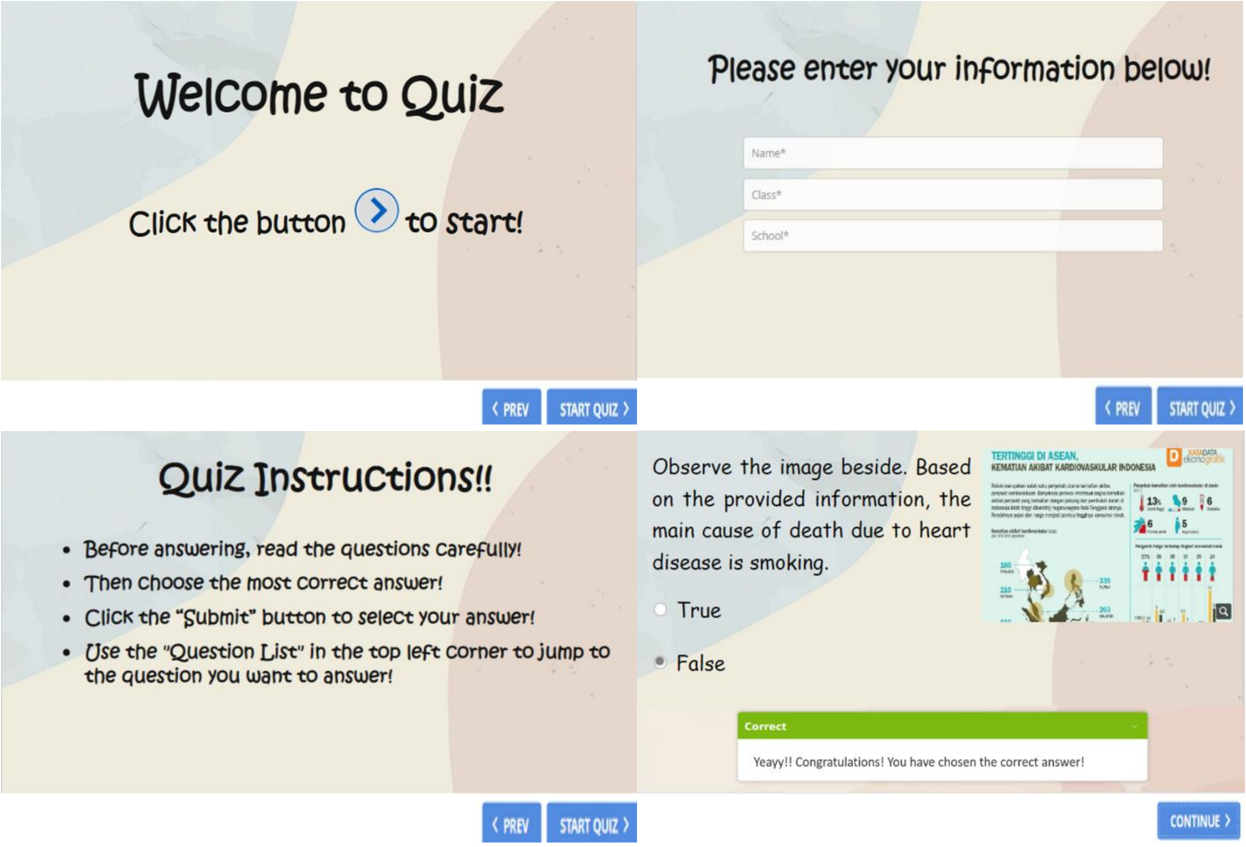


Figure 4b. Quiz Display in English

The process of working on quizzes has automatic time to do. The processing time is equated with the time when working in the form of a paper test. Images in the quiz can also be zoomed in and out to clarify the images provided. Fonts and font sizes are made as straightforward as possible, and the characters are adjusted to the age of students but still clearly legible according to the direction of validators or experts. After completing the work, the results of the student's work are automatically sent to the teacher's email, and immediately, there are scores for multiple-choice questions and short fills. For description questions, teachers provide independent assessments according to student email answers. This process makes it easier for teachers to measure the achievement of learning outcomes. According to Khasanah et al. (2022), measuring science literacy is more accessible and practical because it is combined with e-modules to give students a pleasant learning experience. In line with these findings, Kurniawati et al. (2021) also said that questions to measure competency achievement become more effective when presented in e-modules. However, questions can still be given by paper test or according to the needs and policies of each teacher.

Conclusion

Teachers in elementary schools have not developed many science literacy problems, so the science literacy competence of students has not been measured clearly and accurately, even though science literacy is a competency that students in 21st-century education must possess. This research is a research type of Research and Development (R&D) to develop instruments for measuring science literacy competence. The question items developed consist of 20 items, of which 10 (ten) multiple-choice questions, 5 (five) short fill-in questions, and 5 (five) description questions. The questions developed for assessing science literacy were thoroughly validated through content and empirical methods, confirming their reliability and appropriateness for educational use. Expert evaluations and statistical analyses affirmed that the questions can effectively measure students' understanding of science literacy across various cognitive levels. Moreover, the questions were optimized for digital use, converted into .apk format to ensure accessibility on Android devices, and integrated with e-modules. This digital adaptation not only enhances the accessibility of the assessments but also aligns them with modern educational technologies, making them a valuable tool for measuring science literacy in a contemporary learning environment.

Considering the results of the study, suggestions for future research to develop science literacy problems that contain more indicators of science literacy to measure the competence of science literacy more broadly and comprehensively. The development of science literacy problems that are made digitally, in future research to not only develop questions that can be accessed in the Android program but also those that can be accessed in the iOS program.

Credit Authorship Contribution Statement

Esty Setyo Utaminingsih: Conceptualization, Methodology, Software, Visualization, Formal analysis, Writing – original draft, Writing – review & editing. **Ellianawati**: Conceptualization, Methodology, Formal analysis, Resources, Writing – review & editing, Supervision, **Sri Sumartiningsih**: Visualization, Formal analysis, Funding – review, **FR. Wuriningsih**: Resources, Writing – review, Funding, **Maria Ayu Puspita:** Software, Visualization, Formal analysis, Editing.

References

Adnan, Mulbar, U., Sugiarti, & Bahri, A. (2021). Biology Science Literacy of Junior High School Students in South Sulawesi, Indonesia. *Journal of Physics: Conference Series*, *1752*(1). https://doi.org/10.1088/1742-6596/1752/1/012084

Afriana, J., Permanasari, A., & Fitriani, A. (2016). Project based learning integrated to stem to enhance elementary school's students scientific literacy. *Jurnal Pendidikan IPA Indonesia*, *5*(2), 261–267. https://doi.org/10.15294/jpii.v5i2.5493

Ahied, M., Muharrami, L. K., Fikriyah, A., & Rosidi, I. (2020). Improving students' scientific literacy through distance learning with augmented reality-based multimedia amid the covid-19 pandemic. *Jurnal Pendidikan IPA Indonesia*, *9*(4), 499–511. https://doi.org/10.15294/jpii.v9i4.26123

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings, Educational and Psychological Measurument. *Journal Articles; Reports - Research; Numerical/Quantitative Data*, *45*(1), 131–142.

Amaringga, N. G., Amin, M., & Irawati, M. H. (2021). The effect of problem-based learning module containing research result to improve students' scientific literacy. *AIP Conference Proceedings*, *2330*(March). https://doi.org/10.1063/5.0043529

Arsi, A. (2021). Realibilitas Instrumen Dengan Menggunakan Spss. *Validitas Realibilitas Instrumen Dengan Menggunakan Spss*, 1–8. https://osf.io/m3qxs

Astuti, A. T., Supahar, Mundilarto, & Istiyono, E. (2020). Development of assessment instruments to measure problem solving skills in senior high school. *Journal of Physics: Conference Series*, *1440*(1). https://doi.org/10.1088/1742-6596/1440/1/012063

Bahri, S., Simbolon, M., & Alhad, K. (2021). Development of Assessment Instrument of Scientific Literacy Ability for Students at Musamus University. *Advances in Social Science, Education and Humanities Research Proceedings of the International Joined Conference on Social Science (ICSS 2021)*, *603*(Icss), 395–398.

Bartimote-Aufflick, K., Bridgeman, A., Walker, R., Sharma, M., & Smith, L. (2016). The study, evaluation, and improvement of university student self-efficacy. *Studies in Higher Education*, *41*(11), 1918–1942. https://doi.org/10.1080/03075079.2014.999319

Bashooir, K., & Supahar. (2018). Validitas dan reliabilitas instrumen asesmen kinerja literasi sains pelajaran Fisika berbasis STEM. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *22*(2), 168–181. https://doi.org/10.21831/pep.v22i2.20270

Bucchi, M., & Saracino, B. (2016). “Visual Science Literacy”: Images and Public Understanding of Science in the Digital Age. *Science Communication*, *38*(6), 812–819. https://doi.org/10.1177/1075547016677833

Chung, C. C., Huang, S. L., Cheng, Y. M., & Lou, S. J. (2022). Using an iSTEAM project-based learning model for technology senior high school students: Design, development, and evaluation. In *International Journal of Technology and Design Education* (Vol. 32, Issue 2). Springer Netherlands. https://doi.org/10.1007/s10798-020-09643-5

Ding, C. (2022). Examining the context of better science literacy outcomes among U.S. schools using visual analytics: A machine learning approach. *International Journal of Educational Research Open*, *3*(April), 100191. https://doi.org/10.1016/j.ijedro.2022.100191

Fakhriyah, F., Masfuah, S., Roysa, M., Rusilowati, A., & Rahayu, E. S. (2017). Student’s science literacy in the aspect of content science? *Jurnal Pendidikan IPA Indonesia*, *6*(1), 81–87. https://doi.org/10.15294/jpii.v6i1.7245

Febriyanti, D. F., & Sari, P. M. (2022). Pengembangan Media Pembelajaran Berbasis Literasi Sains Menggunakan Software Ispring Suite 9 pada Pembelajaran IPA. *Jurnal Basicedu*, *6*(4), 6620–6629. https://doi.org/10.31004/basicedu.v6i4.3323

Flores, C. (2018). Problem-based science, a constructionist approach to science literacy in middle school. *International Journal of Child-Computer Interaction*, *16*, 25–30. https://doi.org/10.1016/j.ijcci.2017.11.001

Gurses, A., Gunes, K., Barin, T. B., Eroglu, Z., & Cozel, F. S. (2015). Relation Between Pre-Service Chemistry Teachers’ Science Literacy Levels and Their Some Scientific Process Skills. *Procedia - Social and Behavioral Sciences*, *197*(February), 2395–2402. https://doi.org/10.1016/j.sbspro.2015.07.300

Herlina, Sartono, B., & Susetyo, B. (2020). Kajian Validitas Instrumen Pengukuran Skala. *Indonesian Journal of Statistics and Its Applications*, *4*(1), 136–155.

Ilmi, R., Arnawa, I. M., Yerizon, & Bakar, N. N. (2021). Development of an Android-Based for Math E-Module by using Adobe Flash Professional CS6 for Grade X Students of Senior High School. *Journal of Physics: Conference Series*, *1742*(1). https://doi.org/10.1088/1742-6596/1742/1/012026

Janna, N. M., & Herianto. (2021). Artikel Statistik yang Benar. *Jurnal Darul Dakwah Wal-Irsyad (DDI)*, *18210047*, 1–12.

Jiananda, A. (2017). *Pengembangan Instrumen Penilaian Proyek Mind Mapping untuk Mengukur Penguasaan Konsep Fisika dan Keterampilan Berpikir Kreatif Development of Mind Mapping Project Assessment Instruments to Measure Mastery of Physics Concepts and Creative Thinking Skills*. 1–5.

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*(August). https://doi.org/10.1016/j.stueduc.2020.100935

Kasse, F., Ragil, I., Atmojo, W., Maret, U. S., Thinking, C., Solving, P., & Education, J. (2022). *Analisis kecakapan abad 21 melalui literasi sains pada siswa sekolah dasar*. *10*(1), 124–128.

Khasanah, Q., Prasasti, P. A. T., & Kusumawati, N. (2022). Kelayakan E-Modul Berbasis T-PACK Dalam Memberdayakan Literasi Sains Pada Siswa Kelas IV SDN Pilangbango. *Prosiding Konferensi …*, *3*. http://prosiding.unipma.ac.id/index.php/KID/article/view/3162%0Ahttp://prosiding.unipma.ac.id/index.php/KID/article/download/3162/2510

Kurniawati, T. D., Akhdinirwanto, R. W., & Fatmaryanti, S. D. (2021). Pengembangan E-Modul Menggunakan Aplikasi 3D PageFlip Professional Untuk Meningkatkan Kemampuan Literasi Sains Peserta Didik. *Jurnal Inovasi Pendidikan Sains (JIPS)*, *2*(1), 32–41. https://doi.org/10.37729/jips.v2i1.685

Laliyo, L. A. R., Tangio, J. S., Sumintono, B., Jahja, M., & Panigoro, C. (2020). Analytic Approach of Response Pattern of Diagnostic Test Items in Evaluating Students ' Conceptual Understanding of Characteristics of Particle of Matter. *Journal of Batlic Science Education*, *19*(5), 824–841.

Landis, J. R., & Koch, G. G. (1977). Landis amd Koch1977_agreement of categorical data. *Biometrics*, *33*(1), 159–174.

Lawless, K. A., Brown, S. W., Rhoads, C., Lynn, L., Newton, S. D., Brodowiksa, K., Oren, J., Riel, J., Song, S., & Wang, M. (2018). Promoting students' science literacy skills through a simulation of international negotiations: The GlobalEd 2 Project. *Computers in Human Behavior*, *78*, 389–396. https://doi.org/10.1016/j.chb.2017.08.027

Li, L., Shi, J., Wu, D., & Li, H. (2020). Only child, parental educational expectation, self-expectation and science literacy in Zhuang adolescents in China: A serial mediation model. *Children and Youth Services Review*, *115*(May), 105084. https://doi.org/10.1016/j.childyouth.2020.105084

Magdalena, I., Fauziah, S. N., Faziah, S. N., & Nupus, F. S. (2021). Analisis Validitas, Reliabilitas, Tingkat Kesulitan Dan Daya Beda Butir Soal Ujian Akhir Semester Tema 7 Kelas Iii Sdn Karet 1 Sepatan. *BINTANG : Jurnal Pendidikan Dan Sains*, *3*(2), 198–214. https://ejournal.stitpn.ac.id/index.php/bintang

McComas, W. F. (2019). Programme for International Student Assessment (PISA). In *OECD*. https://doi.org/10.1007/978-94-6209-497-0_69

Moleong, L. J. (2017). *Metodelogi Penelitian Kualitatif*. PT Remaja Rosdakarya Offset.

Muthmainnah, & Istiyono, E. (2019). Constructing Reasoning Multiple Choice Test to Measure Bloomian Higher Order Thinking Skills in Physics of XI Grade Students. *Journal of Physics: Conference Series*, *1233*(1). https://doi.org/10.1088/1742-6596/1233/1/012037

Nani, F. (2021). Analisis Tingkat Kesukaran, Daya Pembeda, Dan Efektivitas Pengecoh Soal Pelatihan Kewaspadaan Kegawatdaruratan Maternal Dan Neonatal. *Paedagoria : Jurnal Kajian, Penelitian Dan Pengembangan Kependidikan*, *12*(2), 199–205.

Noroozi, O., & Mulder, M. (2017). Design and evaluation of a digital module with guided peer feedback for student learning biotechnology and molecular life sciences, attitudinal change, and satisfaction. *Biochemistry and Molecular Biology Education*, *45*(1), 31–39. https://doi.org/10.1002/bmb.20981

Novalia, & Syazali, M. (2014). *Olah Data Penelitian Pendidikan*. Aura Publishing.

Payadnya, I. P. A. A., & Jayantika, I. G. A. N. T. (2018). *Panduan Penelitian Eksperimen Beserta Analisis Statistik Dengan SPSS*. Deepublish.

Polit, D. F., & Beck, C. T. (2006). The Content Validity Index: Are You Sure You Know What's Being Reported? Critique and Recommendations. *ResearchinNursing&Health*,

29, 489–497. https://doi.org/10.1002/nur

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Focus on Research Methods Handling Missing Data in Self-Report Measures. *ResearchinNursing&Health*, *30*, 459–467. https://doi.org/10.1002/nur

Pomalato, S. W. D., Ismail, R., Amin Otoni Harefa, Imawan, O. R., Ningsi, B. A., & Wulandar, D. (2021). Instrument test development of mathematics skill on elementary school. *Mathematics and Statistics*, *9*(2), 106–111. https://doi.org/10.13189/ms.2021.090204

Pratiwiningtyas, B. N., Susilaningsih, E., & Sudana, I. M. (2017). Pengembangan Instrumen Penilaian Kognitif untuk Mengukur Literasi Membaca Bahasa Indonesia Berbasis Model Pirls pada Siswa Kelas IV SD. *Journal of Educational Research and Evaluation*, *6*(1), 1–9. http://journal.unnes.ac.id/sju/index.php/jere

Qadar, R., Haryanto, Z., Subagiyo, L., Junus, M., & Syam, M. (2022). Indonesian Science Teachers' Ability to Design Scientific Literacy Test. *International Journal of STEM Education for Sustainability*, *2*(1), 133–139. https://doi.org/10.53889/ijses.v2i1.52

Rofiyadi, Y. A., & Handayani, S. L. (2021). Pengembangan Aplikasi E-Modul Interaktif Berbasis Android Materi Sistem Peredaran Darah Manusia Kelas V Sekolah Dasar. *JPDI (Jurnal Pendidikan Dasar Indonesia)*, *6*(2), 54. https://doi.org/10.26737/jpdi.v6i2.2575

Rohana, I. N., Rusilowati, A., & Khumaedi. (2018). Pengembangan Tes untuk Mengukur Kemampuan Literasi Sains Siswa SMP pada Materi Getaran dan Gelombang. *UPEJ Unnes Physics Education Journal*, *7*(3), 1–10.

Rokhmah, A., Sunarno, W., & Masykuri, M. (2017). Science Literacy Indicators in Optical Instruments of Highschool Physics Textbooks Chapter. *Jurnal Pendidikan Fisika Indonesia*, *13*(1), 19–24. https://doi.org/10.15294/jpfi.v13i1.8391

Rusilowati, A. (2018). Asesmen Literasi Sains: Analisis Karakteristik Instrumen dan Kemampuan Siswa Menggunakan Teori Tes Modern Rasch Model. *Prosiding Seminar Nasional Fisika Universitas Riau Ke-3*, *September*, 2–15. https://snf.fmipa.unri.ac.id/wp-content/uploads/2019/03/0.-300B-2-15NI.pdf

Schleicher, A. (2018). PISA 2018: Insight and Interpretations. In *OECD*.

Septiani, A. D., Sjaifuddin, S., & Berlian, L. (2022). Pengembangan Instrumen Evaluasi Tes Two-Tier Multiple Choice Berbasis Literasi Sains Siswa Kelas VII Pada Tema Hujan Asam. *Biodik*, *8*(1), 167–174. https://doi.org/10.22437/bio.v8i1.17305

Solichin, M. (2017). Analisis Daya Beda Soal, Taraf Kesukaran, Validitas Butir Tes, Interpretasi Hasil Tes dan Validitas Ramalan dalam Evaluasi Pendidikan. *Dirasat: Jurnal Manajemen & Pendidikan Islam*, *2*(2), 192–213. www.depdiknas.go.id/evaluasi-proses-

Sultan, A. Al, Jr, H. H., & Lickteig, D. (2021). Assessing preservice elementary teachers' conceptual understanding of scientific literacy. *Teaching and Teacher Education*, *102*, 103327. https://doi.org/10.1016/j.tate.2021.103327

Suparya, I. K., I Wayan Suastra, & Putu Arnyana, I. B. (2022). Rendahnya Literasi Sains: Faktor Penyebab Dan Alternatif Solusinya. *Jurnal Ilmiah Pendidikan Citra Bakti*, *9*(1), 153–166. https://doi.org/10.38048/jipcb.v9i1.580

Susanto, H., Rinaldi, A., & Novalia. (2015). Analisis Validitas Reabilitas Tingkat Kesukaran dan Daya Beda pada Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Matematika. *The Journal of the Japan Society for Respiratory Endoscopy*, *37*(3), 343.

Trémolière, B., & Djeriouat, H. (2021). Exploring the roles of analytic cognitive style, climate science literacy, illusion of knowledge, and political orientation in climate change skepticism. *Journal of Environmental Psychology*, *74*(January). https://doi.org/10.1016/j.jenvp.2021.101561

Ulva, T. S., Samitra, D., & Kusnanto, R. A. B. (2021). Pengembangan Instrumen Tes Literasi Sains pada Tema 8 Lingkungan Sahabat Kita untuk Siswa Sekolah Dasar. *Jurnal*

Basicedu, *6*(4), 7174–7187. https://doi.org/10.31004/basicedu.v5i4.1230

Utami, R. T., & Desstya, A. (2021). Analisis Cakupan Literasi Sains dalam Buku Siswa Kelas V Tema 4 Karya Ari Subekti di Sekolah Dasar. *Jurnal Basicedu*, *5*(6), 5001–5013. https://doi.org/10.31004/basicedu.v5i6.1556

Utaminingsih, E. S. (2023). *Pengembangan E-Modul Berbasis STEAM Topik Peredaran Darah Manusia untuk Meningkatkan Literasi Sains dan Karakter Pelajar Pancasila*. Universitas Negeri Semarang.

Utaminingsih, E. S., Ellianawati, E., Sumartiningsih, S., & Puspita, M. A. (2023). STEAM Education. *Jurnal Ilmiah Profesi Pendidikan*, *8*(3), 1605–1612.

Utaminingsih, E. S., Ellianawati, E., Widiarti, N., Sumartiningsih, S., & Puspita, M. A. (2023). A Systematic Review : Digital Literacy for Strenghthening Character in Facing the Era of Society 5.0. *Research and Development Journal Of Education*, *9*(2), 638–647.

Utaminingsih, E. S., & Puspita, M. A. (2023). Emotional Intelligence and Its Important Role. *Jurnal Ilmiah Profesi Pendidikan*, *8*(November 2021), 2003–2011.

Utaminingsih, E. S., Raharjo, T. J., & Ellianawati. (2023). Development of an E-module Based on STEAM on the Topic of Human Blood Circulation. *Jurnal Penelitian Pendidikan IPA*, *9*(7), 5333–5340. https://doi.org/10.29303/jppipa.v9i7.3719

UZ, L. Z., Haryono, & Wardani, S. (2019). The Development of Chemical E-Module Based on Problem of Learning to Improve The Concept of Student Understanding Article Info. *Innovative Journal of Curriculum and Educational Technology*, *8*(2), 59–66. https://journal.unnes.ac.id/sju/index.php/ujet/article/view/31340