Contents list available at Directory of Open Access Journals (DOAJ)
Al-Ishlah: Jurnal Pendidikan
ISSN: 2087-9490 (Printed); 2597-940X (Online)
Journal Homepage:
http://www.journal.staihubbulwathan.id/index.php/alishlah

Development of Authentic Assessment Instrument Based on Scientific Literacy to Measure Science Process Skills

Faizatul Mabruroh[1], **M. Hasbi Ashsiddiqi**[2]
DOI: **10.35445/alishlah.v13i3.287**

Article Info

Keywords:
Instrument;
Authentic assessment;
Science process skills.

Abstract

This research aims to develop an authentic assessment instrument based on scientific literacy to measure students' science process skills. Developing such instruments is significant as it ensures that assessments are more reflective of real-world scientific understanding and practices, potentially leading to improved educational practices. The development stages of this instrument use the research and development (R&D) method with the 4D development model proposed by Sivasailam Thiagarajan, Dorothy S. Semmel, and Melyvn I. Semmel. The procedures for developing this instrument include four stages: definition, design, development, and dissemination. The research subjects were sixth-semester students in the Physics Education Study Program at UIN Raden Fatah Palembang who had taken Basic Physics and Thermodynamics courses. Using a purposive sampling technique, 15 students were involved, divided into three groups (low, medium, high) with five students in each group. The instruments developed were categorized as valid by experts, with a validity percentage of 90% for observation sheets and rubrics, 88% for performance assessment instruments, and 72% for authentic assessment experts. The response test was categorized as practical, with a percentage of 91%.

Kata kunci:
Instrumen;
Asesmen autentik;
Keterampilan proses sains.

Abstrak

Penelitian ini bertujuan untuk mengembangkan instrumen asesmen otentik berbasis literasi sains untuk mengukur keterampilan proses sains mahasiswa. Tahapan pengembangan instrumen ini menggunakan model pengembangan 4D yang dikemukakan oleh Sivasailam Thiagarajan, Dorothy S. Semmel dan Melyvn I. Semmel. Prosedur pengembangan instrumen asesmen ini meliputi empat (4) tahapan yaitu pendefinisian, perancangan, pengembangan, dan penyebaran. Subjek penelitiannya adalah mahasiswa semester VI (enam) Program Studi Pendidikan Fisika UIN Raden Fatah Palembang yang telah memperoleh mata kuliah Fisika Dasar dan Termodinamika. Melalui teknik purposive sampling peneliti melibatkan 15 orang mahasiswa yang dibagi kedalam tiga kelompok (rendah, sedang, tinggi) masing-masing terdiri dari 5 orang. Instrumen yang dikembangkan dikategorikan valid oleh ahli dengan persentase 90% untuk lembar observasi dan rubrik, 88% untuk instrumen asesmen kinerja, dan 72% untuk ahli asesmen otentik. Sedangkan uji respon yang diperoleh dikategorikan praktis dengan persentase 91%.

[1] Universitas Islam Negeri Raden Fatah Palembang, Indonesia
Email: faizatul.mabruroh_uin@radenfatah.ac.id
[2] Universitas Islam Negeri Raden Fatah Palembang, Indonesia
Email: mhasbiashsiddiqi_uin@radenfatah.ac.id

Vol. 13 (3); December, 2021
Received November 28, 2020; Received in revised form: December 3, 2020; Accepted September 25, 2021; Available online December 31, 2021

INTRODUCTION

Ministerial Regulation Number 23 of 2016 concerning Educational Assessment Standards states that one of the principles of assessment is integration, which means that assessment is an inseparable component of learning activities. This principle emphasizes that assessment should be woven into the fabric of the educational process, enhancing and reflecting the learning journey rather than being a detached endpoint. In relation to learning activities, Norman E. Gronlund (1976) formulates the notion of evaluation as a systematic process to determine or make decisions regarding the extent to which learning goals have been achieved by students. This systematic approach ensures that evaluations are comprehensive, objective, and aligned with the instructional objectives. Integrating assessment with learning activities not only supports continuous feedback and improvement but also helps identify individual student needs, thereby fostering a more personalized learning experience. Such a cohesive approach to assessment can significantly enhance educational practices by ensuring that assessments are meaningful and directly tied to the learning outcomes.

Minister of Education of Manitoba (Mujadalah, 2016) also confirms that assessment is a unified whole in science learning and has the main role of getting information related to what is known, what is able to be done, and what students learn. Based on this description, it can be concluded that the assessment is carried out in an integrated manner with the learning process so that the assessment can be used as feedback, to direct learning, and to evaluate students and learning that has been done.

Learning in the 21st century emphasizes students play a more active role in the learning process so that the assessment also experiences a shift, namely from the assessment of the results to an assessment of processes that consider attitudes, behavior, and morals as an inseparable part when assessing learning outcomes. One form of appropriate assessment that meets the current learning measurement needs is authentic assessment. Authentic assessment (authentic assessment) has a strong relevance to the scientific approach in learning in accordance with the demands of the ongoing curriculum.

Authentic-based evaluation aims to help students develop scientific thinking and high-level cognitive skills. In the 2013 curriculum, scientific process skills and high-level thinking skills are crucial components. Stiggins (1994) explains that there are four assessment methods applicable in science learning: Selected Response, Essay Assessment, Performance Assessment, and Personal Communication. The five assessment targets include knowledge, reasoning, skills, product/work, and affective attributes. The implementation of authentic assessment in science aligns with the competency standards of graduates from the 2013 curriculum, which require authentic assessment practices. These practices help students become proficient in demonstrating tasks that mirror real-world scenarios, fostering meaningful engagement and application of knowledge. Ultimately, this approach aims to cultivate students who are not only knowledgeable but also capable of performing practical and relevant tasks, thereby becoming productive citizens.

Science process skills are one aspect of science. Science process skills are used to build the body of knowledge, a collection of science. In addition to scientific process skills, high-level thinking skills that are critical thinking need to be owned by students because it can help students to develop their intellectual potential, as explained by Johnson (2007) which states that the ability to think critical is needed for someone, this is considered important because it is capable Helping students develop their intellectual potential, have the ability to evaluate systematically, and be able to argue in an organized manner.

Performance assessment is considered as an effort to integrate the activities of measuring learning outcomes with the entire learning process, even the assessment itself is an inseparable part of the whole learning process. The main purpose of the performance assessment is to evaluate the actual process in conducting science (Slater, 1993). In addition, the learning objectives of other performance assessments are to be able to develop capabilities (Slater,

1993): 1) Scientific methods, 2) apply systematic procedures, 3) Using materials, laboratory equipment, and computers, 4) Applications and evaluation of various approaches, 5) solve the problem.

In its implementation, especially physics learning is still often found evaluation questions on formative assessments tend to demand memory skills and solve problems through mathematical solutions in the form of multiple choice questions that are often contracted with authentic assessments. The form of multiple choice questions was chosen because this model makes it easy for teachers to make distribution of questions related to the material learned, but it also makes it easier for teachers to conduct assessments because in the form of multiple choice questions, there is only one correct answer. However, the form of multiple-choice tests requires strict supervision so that there is no tendency for students to cheat. Besides that, the multiple-choice form tends to be ineffective when used as an evaluation tool to measure high - level thinking skills in students.

Therefore, in this study, researchers intend to conduct research on the development of authentic assessment instruments in the form of written tests in the form of description questions that can be used to measure student science process skills in thermodynamics material. Written tests in the form of essays/descriptions selected as explained by Nitko & Brookhart (2011) states that the essay test provides more opportunities for students to display the ability to write, organize, express, and explain the relationship between ideas so that they are able to assess the ability of students' high -level thinking skills.

The teaching material to be selected in the development of this instrument is the subject of thermodynamics, because based on the results of interviews in the field, information is obtained that thermodynamics material is one of the essential materials that must be mastered by prospective physics education teachers. Based on the results of observations it is known that in the thermodynamics lecture, the authentic assessment instrument has not been used to determine the ability of the process of science of prospective teachers. So, it is necessary to develop authentic assessment instruments that can measure the skills of students in the science process as candidates for physics teachers.

The research objectives to be achieved in this study are to produce authentic assessment instruments based on scientific literacy to measure the ability of the science process and get an overview of the process of developing authentic assessment instruments based on science literacy developed on aspects of validity and their reliability also obtain information related to the practicality of authentic assessment instruments based on literacy Science in Measuring Science Process Skills.

METHOD

The type of research used is the research and development model (Research and Development) adopted the development model by Sivasialam Thiagarajan, Dorothy S. Semmel and Melyvn I. Semmel (1974), namely the 4-D (Four-D) model with the research stage namely, Define, Design, Develop, and Dissemination.

The description of each of these stages is carried out as follows:

1. Definical Stage (Define). This stage aims to determine and define the learning requirements. What is done at this stage is to analyze to establish basic problems in learning and criteria that are used as a reference in developing assessment instruments. The defining stage is divided into four steps namely initial analysis, student analysis, concept analysis, and specifications of instructional goals.
2. Design Phase (Design). At the design stage, the researcher prepares the prototype of learning tools needed in the research process. The device and support the required assessment instrument consists of four stages, namely the preparation of a benchmark reference test, media selection, selection of assessment formats, and initial design. The preparation of the benchmark reference test is the initial stage to connect the independence

stage with the design stage. This is done by establishing a research instrument grid in accordance with the learning objectives that have been set at the defining stage.
During the implementation of learning activities, learning media is needed with practicum methods so that student worksheets (LKM) or worksheets can be used as guidelines for practicum activities. Furthermore, the preparation of the assessment instrument format is intended to adjust the appearance of the instrument used with the assisted learning media. At this stage the validation and response questionnaire and test questions are also carried out to measure the scientific process skills that will be assessed based on the aspects to be measured and the determination of the questionnaire grid and the assessment rubric.

3. Development Phase (Develop). The development stage produces a revised learning tool through several stages and is validated based on input from experts and product practicality based on user response after a limited trial. Through filling out the product validation sheet the developed assessment of the initial prototype provisions produced at the design stage. After that it is revised and tested to get maximum product results.
The implementation of the trial was carried out in two stages, namely the limited field trial in the form of the instrument readability test and the area test. The instrument readability test aims to check whether the sentence used has been understood by the reader and understood the same as what the author wants (Azwar, 2014: 76). In the readability test students are asked to read and check each instrument item and provide responses and suggestions related to the instrument items made. The results obtained from the readability test of the problem are then used to improve the test instrument items so that a good instrument item is obtained for the next trial, namely the area test.
Data on the results of the readability test will later be used to improve the assessment instrument items before being used for the area test. The test data were analyzed to test the empirical validity and reliability of the instrument. Extensive test results are also used to analyze the achievement of science process skills from students as a sample in this study.
4. Disseminate (Disseminate). This stage is carried out when the product that has been revised at the development phase is then implemented on the actual target after that is done more printing and disseminated so that it can be understood by others and used by educators in the learning process.
This study was conducted in the even semester of the 2019/2020 school year in semester VI (six) students of the Physics Education Study Program UIN Raden Fatah Palembang who had obtained basic and thermodynamics teaching materials. The research subjects were taken from semester VI students, namely as many as 15 people. Researchers conducted an assessment instrument test using purposive sampling technique which is divided into three groups consisting of 5 people. The students selected representing each level of student grouping consisting of high, medium and low levels.

Data analysis is a method used to manage the data collected, while the data analysis technique of the user and response of the user in this study is as follows:

1. Validation analysis. This analysis was carried out by the expert on the research instrument used, namely the Validation Questionnaire of the Observation and Rubric Sheet, the Validation Questionnaire Expert The Performance Assessment Instrument, and Validation Questionnaire Authentic Assessment Based on Science Literacy. The results of the initial product validation are used as a reference for product revision. After the revision is carried out and declared valid by the expert, the product can be used in small group trials.
Data processing is obtained by calculating the number of check lists given in the answer column on the questionnaire with a scale of 1 to 5. The formula used to process validator data is as follows:

$$P = \frac{\sum x}{\sum x_i} x\ 100\%$$

Information:
P = percentage
$\sum x$ = the total number of validator answers
$\sum x_i$ = the total number of ideal values

The results of the percentage of data obtained are converted into quality statements by looking at the criteria in Table 1 below:

Table 1. Criteria for valid level

No	Percentage	Category
1	80% - 100%	Very valid
2	60% - 79,99%	Valid
3	40% - 59,99%	Quite valid
4	20% - 39,99%	Less valid
5	0% - 19,99%	Invalid

(Arikunto, 2016)

2. Practical Analysis. To see the practicality of the product developed, data analysis is used based on the Likert scale that can be seen in Table 2. The percentage formula used according to Sudjiono (2010) is as follows:

$$P = \frac{F}{N}\ x\ 100\%$$

Information:
P = The mean we are looking for
F = the number of existing scores (values)
N = *Number of cases (the number of scores themselves)*

The results of the percentage of data obtained are converted into quality statements by looking at the criteria in Table 2 below:

Table 2. Criteria for Practicality Categories

No	Persentase	Kategori
1	80% - 100%	Very practical
2	60% - 79,99%	Practical
3	40% - 59,99%	Quite practical
4	20% - 39,99%	Less practical
5	0% - 19,99%	Not practical

(Sudjiono, 2010)

FINDINGS AND DISCUSSION

Research and development that has been carried out has produced a product in the form of authentic assessment instruments based on science literacy that can be used in evaluating physics performance, especially on thermodynamic material. The first step taken by the researcher is to carry out the design and development of products in the form of authentic assessment instruments based on science literacy on thermodynamic material. Furthermore, the product that has been developed is validated by three experts in this case is a lecturer in the Physics Education Study Program. After the validation results are obtained, the next step is to revise in accordance with the suggestions and comments given.

After the revision process is carried out the next authentic assessment instrument based on scientific literacy that has been improved is tested in a limited manner in order to obtain the user's response from the product that has been developed. In the early stages of the study, the researcher had carried out a series of design and development that was adjusted to the Sivasialam Thiagarajan, Dorothy S. Semmel and Melyvn I. Semmel (1974) As the name implies, the research and development stage carried out has four stages. It's just that in this study the

researcher only performed 3 stages namely defining (define), design, and development (develop) due to the limitations of researchers both in terms of time, ability, and costs.

Based on the stages of the research that has been carried out, obtained data in the form of interviews with teachers and lecturers and students, validator suggestions and revisions so that they can be tested during the development process and user response to produce valid and practical final products. The results of research at each stage of product development are described as follows:

1. *Define Stage*

At this stage, the researcher analyzes the basic things that occur in learning, especially in thermodynamic material. Through the results of the analysis, it was concluded that an instrument could be needed that could assess the performance process that had a reference based on the expected learning achievement criteria.

This stage consists of four main activities, namely initial analysis, student analysis, concept analysis, and specifications of instructional goals. The stages carried out are as follows:

a. Initial analysis

Initial analysis was carried out through interviews with several parties in this case teachers and lecturers who teach thermodynamics material both at school and on campus. Based on the results of the interviews conducted information was obtained that during the learning process with the practicum method it is not uncommon for educators to conduct group assessments. This is emphasized by evidence of assessment that only sees the results of the group's work. Even though every individual in a group also has different roles and abilities. In addition this also occurs because in one practicum activity the number of students participating in the activity is not small. So that educators can only rely on memory when conducting assessments. These limitations are the obstacles for educators to conduct appropriate and accurate assessments because the assessment is not carried out when students show their abilities. In addition, the performance appraisal of most educators still uses simple assessment instruments because it does not contain detailed and complete psychomotor aspects skills. This is what makes most of the assessments that occur only subjectively done.

Daryanto (2014) states that teachers or teaching managers conduct an assessment with the intention of seeing whether the efforts carried out through teaching have reached the goal. Based on this statement, it can be concluded that by conducting an objective assessment will be able to see the extent to which the learning objectives have been achieved. So as to overcome the problems that occur, it is felt necessary to develop a performance assessment instrument that can evaluate the achievement of learning more accurately and objectively.

Based on the results of the interview also obtained information that thermodynamics material is one of the materials that do a lot of practicum activities. However, instruments to measure students' skills objectively and in more detail are not yet available. The following are examples of assessment sheets used in student performance activities in the laboratory are presented in Figure 2 below:

The assessment sheet Skills Practice

Student's name :
Class :

No	Skills assessed	4	3	2	1	Information

Information:
1 = Not Good
2 = Fair Good
3 = Good
4 = Very Good

Figure 2. Performance Assessment Sheet

b. Analysis of students

Based on the literature study of student analysis activities, results are obtained regarding the characteristics and abilities of students. According to Syah (2013), states that students' motor skills will continue to increase diversity, balance, and strength when he occupies high school and will continue to develop based on the experience he gained. At that age the movement that requires the coordination of a series can be carried out appropriately. The practicum activities applied in learning activities are certainly very helpful for students to show and develop their cognitive and psychomotor abilities. The results of the analysis showed the results that the practicum method was able to increase the enthusiasm of students in learning. In its implementation, practicum activities are divided into stages of preparation, implementation, and closing. After practicum, one of the abilities most commonly gained by students is the ability to understand the material and is more skilled in using tools in the laboratory.

c. Concept analysis

Based on the results of literature studies in concept analysis activities, students' psychomotor skills are obtained when practicum activities are laboratory work skills which include manipulative skills and procedure skills. According to Sofyan (2006), it states that manipulative skills are the skills to use laboratory equipment, especially measuring devices, while procedural skills are the skills to do work tools in a certain order.

The psychomotor skills are then grouped into three stages of practicum activities based on the results of the analysis of students, namely the opening, implementation, and closing activities of the practicum. The results are displayed in the form of a chart as follows:

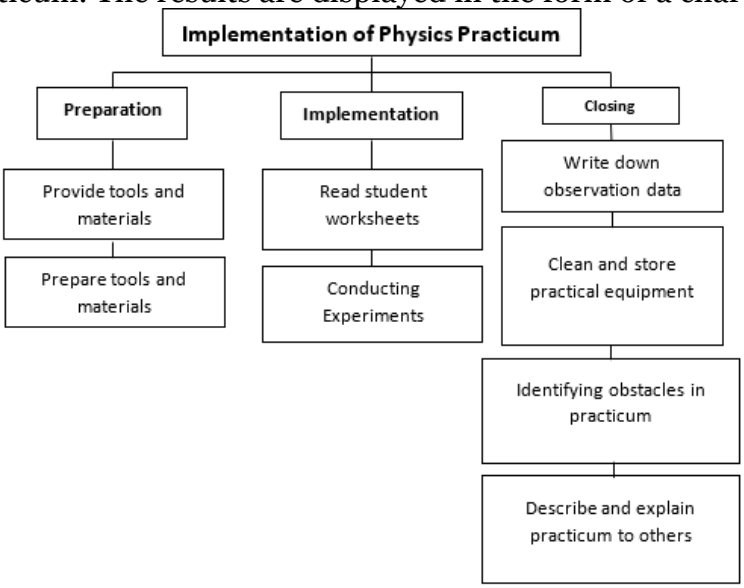


Figure 3. Task Analysis at the time of the practicum

2. *Design stage*

This stage is carried out to prepare the initial design of authentic assessment instruments based on science literacy. This design phase is divided into four steps of activity, namely making an instrument grid, media selection, format selection, and initial design. The complete description of the four steps is as follows:

a. Instrument lattice making

In this step, an authentic literacy-based assessment instrument lattice is produced in thermodynamic material based on the results of previous analysis. The format of the assessment lattice consists of aspects of skills assessment, activities in experimental, psychomotor domains, assessment indicators, item numbers, and number of assessment items. The instrument grids that have been arranged can be seen in table 3 below:

Table 3. Initial Design of Authentic Assessment Instrument Based on Science Literacy

Aspect Evaluation Skills	Activity in Test	Psychomotor Domain	Indicator Evaluation	Number Item	Amount Item
Process Performance	Preparation	Imitation	Prepare tools and materials used	1, 2	2
	Implementation	Manipulating	Use tools and materials	3, 4, 5, 6, 7, 8, 9	7
			Do measurement	4, 7, 10	3
			Clean tools and table Work	13, 14, 15	3
Product Performance	Implementation	Precision	Write down the result data observation	11, 12	2
	Closing	Precision	Convey analysis of result data test	16	1
			Convey conclusion test	16	1
		Articulation	Answer question posttest	17	1

b. Media selection

Authentic Assessment Instruments Based on Literacy Science developed must be supported by appropriate media and in accordance with their use. The media chosen to support this instrument is the practicum worksheet and the tools and materials used. Worksheet is used because it contains questions in the form of practical procedures containing a number of work orders that require psychomotor skills from students

The initial page of the practicum worksheet consists of experimental titles, learning instructions, keywords, while the next page consists of learning objectives based on experiments, basic theory, tools and materials, ways of working, observations, conclusions, constraints in practicum, presentation of experimental results.

c. Format selection

In this step an authentic literacy assessment instrument format is produced in the thermodynamic material which is divided into three stages of activity, namely opening, implementing, and closing. Furthermore, the outline of the Outline of Authentic Assessment Instruments based on Science Literacy by designing the display of the assessment instrument consisting of criteria in the form of aspects of assessment in the form of observation sheets and quality gradation sheets and assessment rubrics. The observation sheet contains the identity of the students, the assessment aspects are divided into three practicum stages, and the column for the score given. While the quality gradation sheet contains aspects of the assessment, value/score, and is equipped with pictures for the preparation and practicum implementation activities.

d. Initial design

Based on the making of the instrument grid, media selection and selection of established formats, then the assessment instrument and the worksheet of students who have been put together. The following is the appearance of the instrument that has been revised:

AUTHENTIC ASSESSMENT INSTRUMENT FOR THE GAS MEETING EXPERIMENT

Student name :

1.

2.

3.

Class :

Date :

Evaluator :

Assessment rubric:

No	Rated aspect	Score
	Preparatory Activities	
1.	Carrying trays/containers with two hands and no tools falling.	4
	Carrying a tray/container with two hands but a tool falls.	3
	Carrying trays/containers with one hand and no utensils falling.	2
	Carrying a tray/container with one hand and the sound of equipment falling.	1
	Do not bring a tray/container containing the tools used in the practicum.	0
2.	Arrange the tools and materials used in the experiment on the work table neatly and check their completeness.	4
	Arrange the tools and materials used in the experiment on the work table neatly but do not check their completeness.	3
	Arrange the tools and materials used in the experiment on the work table neatly and check their completeness.	2
	Arranging the tools and materials used in the experiment on the work table in an untidy manner and not checking their completeness.	1

Figure 4. Authentic Assessment Instrument

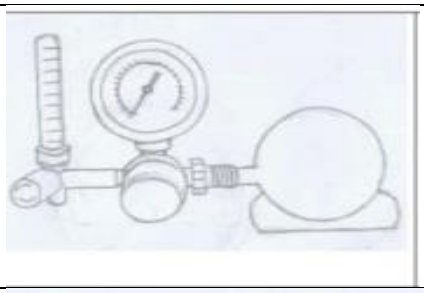
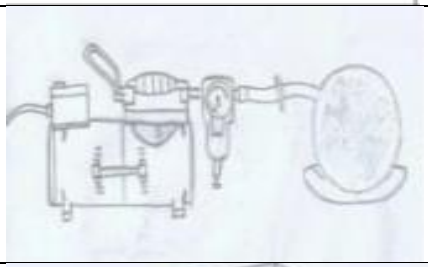
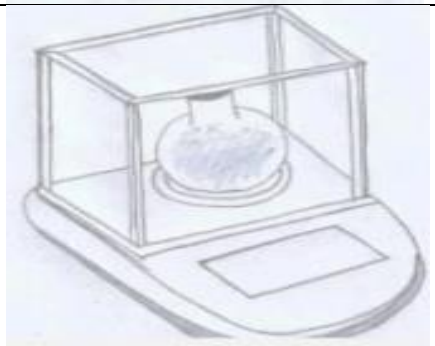
5.	Fill the glass ball with dry air using a barometer at approximately 1 atm					
6.	The glass ball that has been filled with dry air is vacuumed.					
7.	Weigh a glass ball that has been filled with dry air and vacuumed.					

Figure 5. Authentic Assessment Instruments Using Image Instructions

3. Develop Stage

This development stage is divided into four steps, namely validation, revision I, small group test, and revision II. At this stage also will ultimately result in the product development

of authentic assessment instruments based on scientific literassi on valid and practical thermodynamic materials. The explanation of the four steps is as follows:

a. Validation

Product validation is carried out by three expert judgment consisting of authentic assessment validators, observation sheets and rubric validators, and validators of performance instruments. The results of the percentage and validation category are presented in the following table:

Table 4. Validator Assessment Results

No.	Validator Name	Percentage Result	Category
1	Validator of authentic assessment instruments based on scientific literacy	90%	Very valid
2	Validator of observation sheets and rubrics	88%	Very valid
3	Worksheet validator	72%	valid

The results of the percentage of validation of authentic assessment instruments based on science literacy are presented in Figure 6, following:

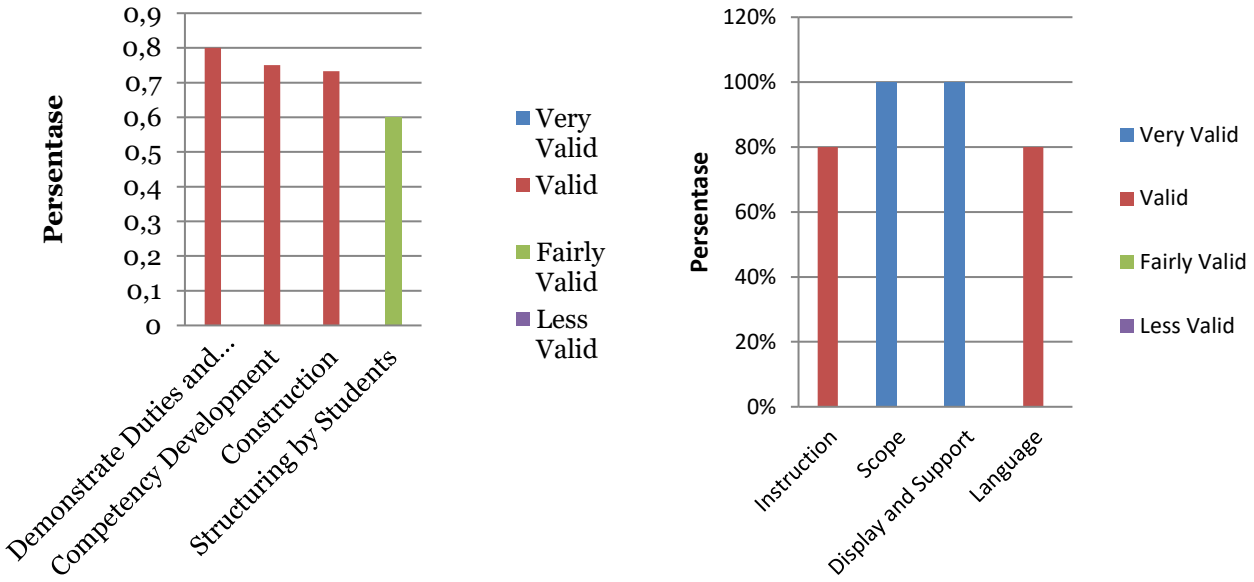


Figure 6. Instrument Expert Validation

Validation questionnaire Observation and rubric sheets consist of four aspects of assessment, namely aspects of instructions, aspects of coverage, aspects of presentation and support, and language aspects. After filling the questionnaire, an average percentage of 88% was produced with the "Very Valid" category. Validation questionnaire Observation and rubric sheet instruments consist of three aspects, namely aspects of content, construction aspects, and language aspects. In the content aspect, the results obtained 87%, the construction aspect obtained 87%results, the language aspect obtained 90%results, and the authentic aspect obtained 88%results. The results of the percentage of Validation Expert Observation and rubrics are presented in Figure 7, following:

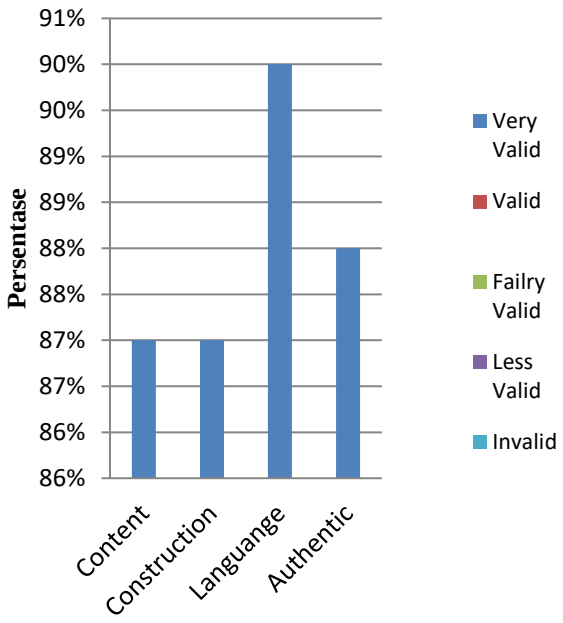


Figure 7. Results of the percentage of Validation Expert Observation and rubric sheets

b. Small group test

After the product of developing an authentic assessment instrument based on a valid scientific literacy followed by a field trial on a small scale. The small group trial was conducted with the subject of research, namely 15 people with division of 3 groups, each group consisting of 5 people who had high, medium and low understanding. The recapitulation of student assessment can be seen in Table 4:

Table 5. Assessment Results Using Authentic Assessment Instruments

Test Gas Meeting							
No	Participant Name Educate	Max Score	Aspect Evaluation			Total Score	Final score
			Preparation	Implementation	Closing		
1	DA	88	6	34	40	80	91
2	RW	88	6	34	40	80	91
3	A	88	6	35	41	82	93.1
4	m	88	7	36	40	83	94.3
5	YM	88	8	35	40	83	94.3
Ideal Gas Experiment							
6	Y	92	6	34	48	88	95.6
7	m	92	6	34	48	88	95.6
8	AP	92	6	34	48	88	95.6
9	NT	92	6	32	38	76	82.6
10	RP	92	6	29	37	72	78.2
Test Equilibrium and Concepts Temperature							
11	SI	92	8	26	48	82	89.1
12	m	92	8	30	48	86	93.4
13	FD	92	8	30	47	85	92.3
14	AF	92	8	28	46	82	89.1
15	A A	92	8	26	46	80	86.9

The percentage of the results of filling in questionnaires per aspect based on the assessment of eight respondents is displayed in Figure 8 below:

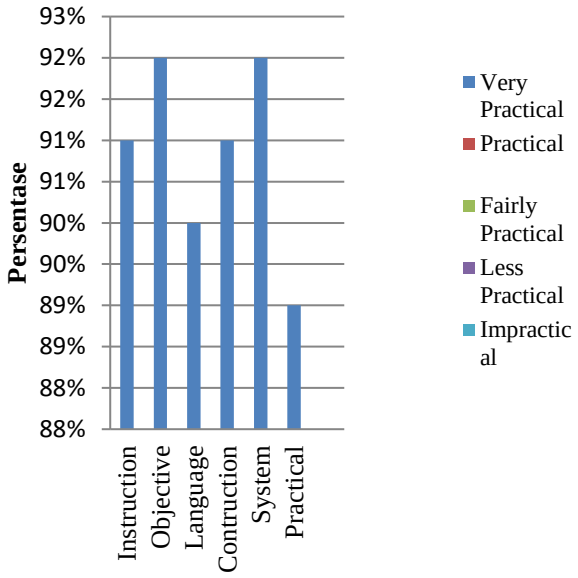


Figure 8. Percentage of the results of filling out the user response questionnaire

The data obtained from the questionnaire indicate that users did not suggest any improvements for the performance assessment instrument. Respondents reported that the instrument is easy to use and allows for objective assessment of students' abilities, particularly in the psychomotor domain. Consequently, it can be concluded that the developed product, based on authentic assessment, is considered "practical" for evaluating heat transfer topics.

Product validation involved three expert validators: those specializing in observation and rubric sheets, performance assessment, and authentic assessment. An instrument is deemed valid if the analysis results meet predetermined criteria. The research results, as previously presented, indicate that the average validation percentages affirm the development of authentic assessment instruments based on scientific literacy as "valid." This finding aligns with Widoyoko (2014), who stated that an instrument is valid if it accurately measures what it is intended to measure. Validity, therefore, pertains to the accuracy of a measuring instrument. Utilizing valid assessment instruments ensures the generation of valid data. Thus, if the data produced by an instrument are valid, the instrument itself can be considered valid.

The instrument's practicality was evaluated during the small group test stage, followed by the distribution of a user response questionnaire. The assessment instrument was used as an evaluation tool in learning activities involving practical methods. According to Arifin (2011), practicality refers to the ease of preparing, using, processing, interpreting, or administering a test. Given the very high average percentage of positive user responses, it can be concluded that the developed authentic assessment instruments based on scientific literacy are "practical."

s

CONCLUSION

The development of Authentic Assessment Instruments based on scientific literacy was carried out through defining and designing stages. This involved identifying standards, selecting media for authentic tasks in student worksheets, choosing appropriate formats, determining assessment criteria, and developing rubrics. The performance assessment instrument developed was validated by expert validators, with the observation and rubric sheet receiving a validity percentage of 90%, the performance assessment instrument receiving 88%, and the authentic assessment experts giving a validity percentage of 72%. Additionally, the

performance assessment instrument was deemed practical based on user responses, achieving a practicality percentage of 91% from eight respondents.

Despite the positive outcomes, this research has several limitations. Firstly, the sample size was relatively small, involving only 15 students from a single institution, which may limit the generalizability of the findings. Secondly, the study focused on specific science subjects (Basic Physics and Thermodynamics), and the applicability of the developed instruments to other scientific disciplines remains uncertain. Finally, the validation process, while thorough, relied heavily on expert judgment, which could introduce subjective biases. Future research should consider expanding the sample size and including participants from diverse educational institutions to enhance the generalizability of the findings. Additionally, it would be beneficial to test the developed instruments across a broader range of scientific disciplines to determine their versatility and effectiveness. Future studies should also explore the integration of technology in authentic assessments to streamline the evaluation process and potentially increase the accuracy and reliability of the assessments. Finally, incorporating longitudinal studies could provide insights into the long-term impact of authentic assessments on students' scientific literacy and process skills development.

REFERENCES

Anjarsari, Putri. (2014). Pentingnya Melatih Keterampilan Berpikir (*Thinking Skills*) dalam Pembelajaran IPA SMP. *Makalah,disampaikan salam PPM "Optimalisasi Implememntasi Kurikulum 2013 dengan Workshop Pengembangan LKS IPA Berpendekatan Guided-Inquiry untuk Mengembangkan Thinking Skills dan Sikap Ilmiah Siswa, tanggal 23 Agustus 2014*. Yogyakarta: Universitas Negeri Yogyakarta.

Arifin, Z. 2011. *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosdakarya

Arikunto, Suharsimi. (2009). *Dasar-dasar Evaluasi Pendidikan (Edisi Revisi)*. Jakarta: Bumi Aksara.

Azwar, Saifuddin. (2003). *Reliabilitas dan Validitas*. Yogyakarta:

Pustaka Pelajar. Azwar, Saifuddin. (2014). *Penyusunan Skala Psikologi*. Yogyakarta: Pustaka Pelajar.

Azim, S., Khan, M. (2012). Authentic assessment: An instructional tool to enhance students learning. *Academic Reasearch International*, 2(3): 314-320. Borich & Kubiszyn. (2013). *Educational testing & measurement: classroom application and practice, (10th ed)*. United States of America: Wiley.

Burhanuddin. (2015). Evaluasi Keterlaksanaan Penilaian Otentik (*AuthenticAssessment*) Pada Pembelajaran IPA SMP Negeri di Pasangkayu Kabupaten Mamuju Utara. *Tesis*. Yogyakarta: PASCA UNY.

Darmayanti, N.W. S., W. Sadia, & A.A.I. A.R. Sudiatmika. 2013. *Pengaruh Model Collaborative Teamwork Learning Terhadap Keterampilan Proses Sains dan Pemahaman Konsep Ditinjau Dari Gaya Kognitif*. e-Journal Program Pascasarjana Universitas Pendidikan Ganesha. Diakses pada 22 Oktober 2019.

Daryanto. 2014. *Evaluasi Pembelajaran*. Jakarta: Rineka Cipta.

Devi, Poppy Kamalia. (2013). *Keterampilan Proses dalam Pembelajaran IPA untuk Guru SMP*. Bandung: PPPPTK IPA.

Dimyati dan Mudjiono. (2002). *Belajar dan Pembelajaran*. Jakarta:

Rineka Cipta. Dimyati dan Mudjiono. (2006). *Belajar dan Pembelajaran*. Jakarta: Rineka Cipta. Djaali & Muljono, P. (2008). *Pengukuran dalam Bidang Pendidikan*. Jakarta: Grasindo.

Ennis, R.H. (2000). *A Super-Steamlined Conception of Critical Thingking*. Diunduh pada 14 Oktober 2017 pukul 14.00 dari http://www.ed.uine.edu/EPS/PESyearbook/92.does/ennis.htm.

Giancoli, D.C. 2014. *Fisika: Prinsip dan Aplikasi*. Jakarta: Erlangga.

Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2004). A Five-Dimensional Framework for Authentic Assessment. *Educational Technology, Research and Development,* 52: 67-86.

Haliday, D. et al. (2010). *Fisika Dasar Edisi Ketujuh Jilid 1*. (Terjemahan Tim Pengajar Fisika ITB). Jakarta: Erlangga. (Edisi asli diterbitkan tahun 2005 oleh John Willey & Sons Inc).

Hambleton, R. K., & Swaminathan, H., (1985). *Item Respon Theory Principles and Applicattion*. Boston: Kluwer. Nijhoff Publishing.

Mabruroh, F, Suhandi, A. (2017). *Construction Of Critical Thinking Skills Test Instrument Related The Concept On Sound Wave Construction Of Critical Thinking Skills Test Instrument Related The Concept On Sound Wave*. https://doi.org/10.1088/1742-6596/755/1/011001

Nitko, A.J., & Brookhart, S.M. (2011). *Educational assessment of student (6th ed)*.Boston: Pearson.

Sofyan, A., Feronika, T., & Milama, B. (2006). *Evaluasi pembelajaran IPA berbasis kompetensi*. Jakarta : Lembaga penelitian UIN Jakarta dan UIN Jakarta press. Mohrig, J.R., Hammond, C.N., & Schatz, P.F. (2010). *Techniques in organic chemistry*. Inggris : W.H. Freeman and Company

Syah, Muhibbin. (2013). Psikologi Belajar. Surabaya: Remaja Pendidikan.

Widoyoko, E.P. (2014). *Evaluasi Program Pembelajaran*. Yogyakarta: Pustaka Pelajar.