



The Development of Test Instruments to Measure the Science Literation Skills of Junior High School Students in Global Warming Themes

Umi Roufatuz Zahro✉, Woro Sumarni, Suharto Linuwih

Pascasarjana, Universitas Negeri Semarang, Indonesia

Article Info

Article History :
Received February 2020
Accepted March 2020
Published April 2021

Keywords:
Scientific literacy, test
instruments, global
warming

Abstract

The progress of science and technology in various countries has accelerated toward the 21 century. The measurement of scientific literacy is very important to find out the extent to which students have been literated in science as an effort to improve the quality of education in Indonesia. The preparation of test instruments based on scientific literacy is an effort to measure the literacy skills of students in the field of science. The purpose of this study was to develop a valid, practical and effective science literacy test instrument to measure the scientific literacy skills of students on the theme of global warming. The research method used in this study was the method of Research and Development (R & D) with the research design of ADDIE (Analysis, Design, Development, Implementation, and Evaluation). The final product in this research development is a scientific literacy-based test instrument to measure the scientific literacy skill of students that have been declared valid and feasible by experts and is valid in the content, in the form of 20 reasonable multiple-choice questions that are tailored to the indicators of the scientific literacy, complete with a grid of questions, answer keys, and bibliography. Based on the results, it is concluded that the validity of the test instruments received an average percentage of 79.99% and was included in the category of valid and feasible to use. The practicality of the test instrument is in the very practical category seen based on the results of student responses by 87.5% and teacher responses by 94.23%. The effectiveness of the test instrument is in the effective category, Since this study can reveal or position the scientific literacy skills of students into three criteria items, namely high, medium, and low, it can be used as input and thought contribution in an effort to improve the quality of making the test instruments to measure the skill of scientific literacy on the theme of global warming.

✉ correspondence :
Pascasarjana, Universitas Negeri Semarang
Jalan Kelud Utara III No. 37, Kota Semarang, Jawa Tengah,
Indonesia 50237
E-mail: zahro.umie@gmail.com

INTRODUCTION

The progress of science and technology in various countries has been growing rapidly toward the 21st century. Science education as one of the subjects in the Junior became an important foundation in the form of qualified human resources. Science education can explain various natural phenomena or phenomena that occur in everyday life. Pendidikan science has an important role in preparing children to enter the world of life.

Science education has great potential and strategic role in preparing qualified human resources to face the era of industrialization and globalization. This potential can be realized if it is able to bring science education of students proficient in their field and managed to cultivate logical thinking skills, creative thinking, problem-solving skills, critical, master technology and adaptive to changes and the times. Thus the process of science education that create a literate man of science (scientific literacy) and the technology completely.

Scientific literacy is considered a key learning outcomes in education at the age of 15 years for all students, for children aged 15 years already should determine career choices and participate take a role in the advancement of science and technology (OECD, 2018). The quality of the education, especially science education in Indonesia is still low when compared with other developing countries. Weak education in Indonesia, especially science education indicated by the low achievement levels of scientific literacy in PISA (Program for International Student Assessment).

Achievement level of scientific literacy Indonesia for 12 years participation has always ranked fifth bottom when scientific literacy is very important in determining the quality of education in a country (Fu'adah et al., 2017). According to data from the Trends in International Mathematics and Science Study (TIMSS) Indonesian students' science literacy scores in 1999, 2003, 2007, 2011, and 2015 are respectively 492, 510, 471, 426, and 397. In 2015, Indonesia was the the order of 44 of 47 participants (Martin, et al., 2016). Meanwhile, according to the Program for International Student Assessment (PISA), Indonesia was ranked 64th out of 65 participants who participated in PISA in 2012

(Nisa, et al., 2015). 2015 Indonesia is progressing with ranks 62nd with a score of 403 out of 70 participants, but far short when compared to Thailand which is ranked 54 with a score of 421 (OECD, 2016). Based on the new PISA report released, Tuesday, December 3rd 2019, the score read Indonesia was ranked 72 out of 77 countries, science score was ranked 70 out of 78 countries (OECD, 2018). The data shows the level of scientific literacy of students in Indonesia is still very low. The level of low scientific literacy of students can be caused by technical instruments that do not accommodate the full criteria for assessing the scientific literacy (Permanasari, 2011). According to research Sudiarmika (2010), a collection of test used in schools are testing this scientific knowledge of the cognitive and mathematical calculations, while the aspects of the process and context escaped assessment.

Interviews with several science teachers SMP in Kudus and analysis of issues prepared by the teacher, there was information that the problems used to measure student learning outcomes have not been charged scientific literacy, the question is more emphasis on the content and not loading aspects of the process and context. Based on the facts in school, which is focused on research issues are the students' science literacy problem, because based on the observation of the questions used in the schools, that question yet included aspects of scientific literacy. Therefore, the measurement of scientific literacy is very important to find out the extent to which the student has meet the science literacy, and thus improving the quality of education in Indonesia can be done and can compete with other countries.

The position of an assessment as an assessment tool is vital in the world of education (Aji, 2015). Preparation of test instruments based on scientific literacy is one way to measure the literacy skills of students, especially in science or science. Sulistiawati (2015) states that to measure students' science literacy skills can be used on a few questions of PISA. Pengembangan measuring devices literacy in PISA involves three aspects: content, process aspects, and aspects of the context. Aspects of science content refers to the key concepts of the science needed to understand natural phenomena and changes made to nature through human activity. The process of science refers to the mental

processes involved when answering a question or solve a problem.

Measurement science literacy is important to know the extent to which literacy students to the scientific concepts they have learned. Competence science students to be low because students are not trained to express opinions or ideas in their mind, so that when given a problem related to the significance and relevance of the material to the surrounding environment students can not afford, (Mardhiyyah et al., 2016). Therefore, we need a literacy instrument sains. Pengukuran scientific literacy is important to know the extent to which literacy students to the scientific concepts they have learned. Therefore, an instrument instrument of scientific literacy is needed.

Instrument scientific literacy although already exist and can be adopted from international studies such as PISA, but the results of scientific literacy Indonesia in international studies applies generally. The diversity of students' backgrounds and curriculum in the educational unit level adapted to the local area as well as the specifications of science subjects, especially science, the authors develop scientific literacy test instruments for use in small and within the scope of science subjects.

Based on this background it is necessary for the development of test instruments to measure the ability of junior high school students in Kudus in science literacy on the theme of global warming.

METHODS

The method used in this research was the method of research and development, or often referred to as the research and development (R & D) with a research design ADDIE (Analysis, Design, Development, Implementation, and

Evaluation) in the adoption of Dick & Carey (Dick & Carey: 1996). The steps of the development is carried out as follows.

Phase analysis (Analysis) in this study is an early stage, which is conducted interviews with relevant science teacher and school observations include SMP N 1 Undaan, MTS Nahdlotul Muslimin, and SMP IT Qolsaba. As for problems found are (1) the instrument test leads students to memorize learning material that has been delivered; (2) The test instrument used is not specifically measure the ability of science literacy of students in learning science theme of global warming; (3) has not developed scientific literacy test instruments theme of global warming. The results of this observation is the basis in determining the design of the test instrument. Developed test instrument design is the development of test instruments to measure students' science literacy skills on the theme of global warming.

Stage design (design) initial test instrument was developed in the form of multiple choice questions grounded as much as 40 eggs, which are adapted to scientific literacy indicators. Questions developed intertwined with reading texts or articles, tables, or images, so that students answer questions based on analysis of reading text, tables, or images that are presented. The initial design of test instruments validated by three experts, namely the expert evaluations, expert materials science and science literacy expert. The initial design of the instrument before being validated contains: a cover page, preface, table of contents, list of aspects and indicators in science literacy, grating matter, manuals, questions, about the scientific literacy test, and a list of questions in the literature of the instrument. Each of the scientific literacy based test that has developed a text relating to the question that the students' can ask in the articles.

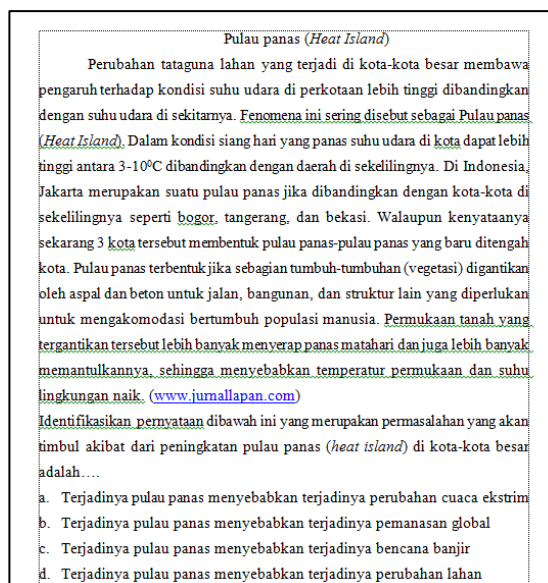


Figure 1. Preliminary Design of Science Literacy Test Instruments

The development phase (development) instrument tests performed after designing the test instrument. Developed test instruments and has passed the stage of improvement include: the cover page, the preface, table of contents, list of aspects and indicators in science literacy, general instructions about the work, about the scientific literacy test, and a bibliography. As a matter of scientific literacy tests that have been developed are multiple choice questions reasoned. Characteristics of matter developed consists of indicators of scientific literacy, problem interlaced with reading texts, articles, pictures, charts, or tables to help students answer about scientific literacy. Aspects developed in the matter is the aspect of content, process aspects, and aspects of attitude. Number of questions developed by 20 MCQs reasoned. According Sumarni et al., (2016) evaluation stage (evaluation) is done by evaluating the results of user testing whether the test instrument capable of measuring students' science literacy skills. Validation of the instrument was tested by calculating the validity of the items and use the questionnaire sheets were tested on three expert lecturers. Expert lecturers selected from subject matter experts, expert evaluation, and expert scientific literacy, which is derived from chemistry and biology lecturer of the State University of Semarang. Small-scale trials conducted to 10 students of grade VII F SMP 1 UndaanKudus.

Small-scale trials aim to determine the assessment or student and teacher responses to the legibility of scientific literacy test instruments that

have been developed, the time required, and the validity of the item that includes distinguishing, level of difficulty, reliability, and quality of posing questions. Large-scale trials conducted to 30 students of grade VII C SMP 1 Undaan Kudus.

RESULTS AND DISCUSSION

The data obtained and analyzed in this research and development used to answer the research problem, so that the research objectives can be achieved. As for the discussion on research and development has been conducted on the level of validity of the instrument, the feasibility of the instrument, and the effectiveness of the instrument.

Test of the validity of the test instruments

The validity of the test instrument in the study include content validity and construct validity. Validity determine the extent to which the instrument could measure in question (Bashoor & Supahar, 2019). Validity of the content used to calculate the level of difficulty of questions, different power, validity, and reliability problems which is calculated using Anates application 4.0.

Furthermore, the test instrument to test the validity in construction by experts. Construction validity was tested using a questionnaire sheets were tested on three expert lecturers. Expert lecturers drawn from the materials experts, expert evaluation, and expert scientific literacy, which is derived from chemistry and biology lecturer of the State University of Semarang.

The validity of the construction work being done includes two stages of testing. Validation Phase 1 study after the development phase and before the small-scale trials. The following is the results of construction validity of the results of the expert validation.

Table 1. Results of the Phase 1 of Validation Expert

Total	Percentage (%)	Information
Validator 1	86.36	Can be used without revision
Validator 2	72.72	Can be used with minor revisions
Validator 3	71.81	Can be used with minor revisions

Design validation phase I test instrument intended for a judgment of an expert on the completeness of the components in the instrument tests, readability, and the contents of the instrument. The results of validation by each expert on the assessment of the first phase, said test instruments can be used without revision if it gets a score of > 75%.

Validation phase I validated instrument to three expert lecturers. The results of each validation of experts to get a score of 86.36%, 72.72% and 71.81%. In addition, the validator also provide input and advice to the test instrument has been developed. The advice given by the validator are as follows: (a) On the cover of the instrument repaired by adding writing the theme of global warming as the theme of the instrument developed, (b) Improvement in the preface changed to the preface and acknowledgments in the preface to be replaced addressed to the party more specific is the validator, counselors, teachers and students who have been involved in the development and refinement of instruments, (c) improvement in attitude aspect that makes the table as another aspect to be more clear, (d) improvements to the grammar matter and punctuation matter and matter are spaced 1 space only, Adding to the picture on the matter.

The results of the validation phase 1 and then analyzed and revised based on the comments or suggestions of experts to be tested on a small scale test. After a small-scale test, the instrument analyzed and corrected item validity of the results of expert validation. Then the results of the repairs back to the expert lecturers validated before tested

in large-scale test besar. Therefore, the expert validation phase 2 results are listed in the following Table 2.

Table 2. Results of Phase II of Validation Expert

Total	Percentage (%)	Information
validator 1	88.18	Can be used without revision
validator 2	76.36	Can be used without revision
validator 3	75.45	Can be used without revision
Average	79.99	Can be used without revision

Based on the validation results of the three experts obtained an average result of 79.99% and a validation of said instrument can be used without revision. So that after the validation phase II instrument can be used for large-scale test.

Small-scale trials conducted after the test instrument repaired based on the input of experts. In small-scale trials are conducted testing this instrument validity test to retrieve data items, retrieving data questionnaire student and teacher feedback questionnaire data. Small-scale trials aimed to determine assessments / student and teacher responses to the legibility of scientific literacy test instruments that have been developed, the time required, and the validity of the item.

Small-scale trials conducted in SMP 1Undaan Kudus. Students sampled in this stage as many as 10 people, selected using purposive sampling techniques, the criteria used are student students with the ability of low, medium, and high based on their value.

Results of student feedback about less time to work on the problems that can be caused by matter and the reading is too long. This causes the students less time in doing. For this, it was revised to a matter that is not too long. The results of data analysis such as the validity of the item as well as guidance to repair the problems. So the problem with different power ugly and invalid question is no longer used for large-scale testing.

Based on the 40 questions that have been made, the validity test items and test the validity of the construction by experts. The validity of the first is to calculate the validity of items that include the level of difficulty, different power, validity, and reliability problems. For the validity of the

construction done is test the instrument test based on suggestions/ improvements of experts ie expert lecturers.

Calculation of the validity of the items do when performing small-scale trials. About the difficulty level of the small-scale trials have criteria difficult, moderate and easy. The results of the 40 questions, the results obtained about the difficulty level as many as 13 problems including difficult criteria, including the criteria for being about 21, and 6 assessment including mudah. Hasil criteria different power calculation, obtained about the criteria of good once as many as 12 questions, questions with good criteria as many as 16 questions, about the criteria pretty much as 7 questions, and the questions with as much as 5 about the ugly criteria.

When distinguishing > 0.40 then, the matter can be used for large-scale testing without revision, if $0.30 < D < 0.39$ and valid question, it can be used for large-scale test without revision, if $0.30 < D < 0.39$ and the question is not valid, it can be used for large-scale test with revisions, if $0.20 < D < 0.29$, it can be used for large-scale test with revisions, and if $-1.00 < D < 0.19$ then the matter can not be used for large-scale test (Sudiatmika, 2010). Obtained distinguishing matter referenced in the decision making items, so the matter with ugly criteria automatically the question is not used.

The analysis of the data obtained a valid question of 26 questions and problems which are not valid of 14 questions. The calculation of reliability question was intended to determine the level of reliability problems, therefore, it is feasible or not the matter if developed and used to measure the students' science literacy skills. Results of the reliability calculation using the formula KR-20 trials of small and large scale. The following are the results of a small-scale test of reliability problems and large-scale test.

Table 3. Results of the Reliability Tests

Reliability	Criteria	
Small Scale Test	0.65	High
Large-scale test	0.73	High

Test instrument reliability is sought from the calculation results of test on a small scale trials and large-scale testing. The result of the calculation using the formula reliability of cognitive problems KR-20 at 0.65. These results suggest that scientific

literacy assessment instrument based multiple-choice type of reasoned otherwise reliable. In a small-scale trial is the matter of the level of reliability of 0.65 and showed a high level of reliable ie. In large-scale testing, the reliability about the show results and can be categorized as high 0.73 so the instrument test developed can be said to be reliable and fit for use.

Quality humbug matter consists of five criteria: excellent, good, less good, bad, and very ugly. Humbug about the already functioning quite well. Humbug in multiple-choice type questions grounded in fact does not affect the quality of the matter, because the score is not only determined by the choice of answers but also determined from the written reasons students (Ridwan et al., 2013). Test instruments that have been declared valid, the instrument of the tests can be used to test the students, while the test instrument which has been declared a reliable means the instrument can be used on the same students in different times, and the result will remain the same, because it has a degree of consistency is high.

Student and teacher given their opinion using a questionnaire to assess the readability test instrument that was developed. Questionnaire responses of students and teachers in this study using a questionnaire with a scale model of the type of rating scale. According to Sugiyono (2012), the raw data from the questionnaire with a scale rating scale models obtained in the form number and then ditafsirkan in a qualitative sense. Data obtained from the questionnaire responses of students and teachers in the form of numbers, and then analyzed and interpreted in percentage. Here is some feedback from the students: (a) there are less obvious image, (b) reading is too long so that the working time is less, (c) writing in the chart is less clear.

The instrument has been tested on a small scale test, got some feedback from students and teachers. Feedback from students and teachers as a reference to correct the instrument tes. Hasil questionnaire responses of students and teachers on a large scale trials show positive results with average results show a rise in every aspect compared to the small-scale test. These results indicate the instrument has been improved so that students and teachers on average memberikn good comment on aspects assessed.

The results of the questionnaire responses of students and teachers indicated that instrument based test scientific literacy that developed has been good, because it has to be improved instrument based test scientific literacy after testing a small scale, whereas the results of the analysis of items showed that 26 question valid criteria, level of difficulty about the difficult, moderate, easy, and distinguishing enough, good, and excellent. A total of 20 questions from 40 questions in the test instrument based on scientific literacy, then used to test usage.

The result of the improvement of small-scale trials then validated again to the experts in order to obtain a valid and worthy instrument. The results of the validation of third-ahli can be seen in Table 4.

Table 4. Summary of Phase II of Validation Expert

Total	Percentage (%)	Information
Validator 1	88.18	Can be used without revision
Validator 2	76.36	Can be used without revision
Validator 3	75.45	Can be used without revision
Average	79.99	Can be used without revision

Based on input from experts, an instrument that has been developed after fixing to get a score of each validator that is 88.18%, 76.36%, and 75.45% and the average obtained is 79.99%. These results indicate that the test instrument was developed declared valid and feasible for Use. According to (Akbar, 2013) declared valid test instruments and can be used without revision if it meets the requirements above 75.00%. Thus, we can conclude the instruments that have been developed through expert validation and twice declared invalid and unfit for use because it meets the criteria of more than 75.00% that get an average score of 79.99%.

The end product in development research is scientific literacy based test instruments to measure students' science literacy skills that have been declared valid and feasible by experts and valid contents. The end product scientific literacy based test instruments in the form of 20 multiple-choice items reasoned adapted to scientific literacy indicators, complete with lattice question, answer key, and bibliography. Therefore, the final product image of science literacy test instruments can be seen in the following figure.



No. Soal	Aspek Literasi Sains			Indikator Literasi Sains
	Kontais	Aspek Konten	Aspek Sikap	
17	Personal	Sistem Bumi	Komunikasi terhadap isu sains	12

37. Amati diagram berikut ini!

(Sumber: Buku Kaya tentang parik)

Diagram tersebut menunjukkan gas karbon dioksida yang digunakan ataupun dilepaskan dalam beberapa proses. Pasangan empat proses yang tepat pada diagram di atas adalah

- 1 - fotosintesis, 2 - respirasi, 3 - pembakaran, 4 - pembusukan
- 1 - respirasi, 2 - pembakaran, 3 - pembusukan, 4 - fotosintesis
- 1 - pembakaran, 2 - pembusukan, 3 - fotosintesis, 4 - respirasi
- 1 - pembusukan, 2 - fotosintesis, 3 - respirasi, 4 - pembakaran

No. Soal	Aspek Literasi Sains			Indikator Literasi Sains
	Kontais	Aspek Konten	Sikap sains	
12	Personal	Literasi Ilmiah	Dialogan terhadap keragaman ideologi	19

Figure 2. Final Design of Literacy Test Instruments

Practicality Test Test Instruments

Practicality test instrument responses obtained from the student and teacher responses from questionnaires given to small-scale trials and

large scale. The average response of teachers and students give positive response to the instruments developed.

Table 5. Summary of the Questionnaire Responses of Students

Student feedback	Percentage (%)	Criteria
Small Scale Test	78	Well
Large-scale test	87.5	Very good

Table 6. Summary of the Questionnaire of Responses Teacher

Student feedback	Percentage (%)	Criteria
Small Scale Test	76.92	Very good
Large-scale test	94.23	Very good

Table 6. Based on the average score student responses by 78% with good criteria to the application of the test instrument on small-scale trials. Suggestions or comments given student is At trial instrument based test large-scale scientific literacy percentage obtained was 87.5% or included in the criteria very well. All aspects of the assessment of scientific literacy based test instruments showed an increase because of improvements have been made in science literacy based test instruments are developed that make improvements to the reference image in question, clarify the picture, fix the problem that is too long.

Teacher response results showed an average of 76.92% on a small scale test and 94.23% on a large scale test. Of the three respondents, teachers respond to the criteria very well. However, an analysis of item questionnaire indicate that there are some things that need to be improved to obtain practical use of the instrument, especially legibility problems and working time problems that must be repaired. After the test instruments improved both in terms of legibility, processing time, and readings in response to the teacher about the large-scale test on average showed an increase of 94.23% and included a very good category. Based kiteria practicality that the device is said to be practical when 80% or more of respondents giving a positive response to a device developed Hobri (2009).

The Effectiveness of the Test Instruments

The calculation of the effectiveness of the test instrument based on the students' science literacy assessment. The results of scientific literacy of students can be categorized into three criteria: low, medium, and high. The level of students' science literacy skills are calculated on usage test conducted at SMP 1 Undaan as many as 75 students. The

results of scientific literacy categories contained in thefigure.

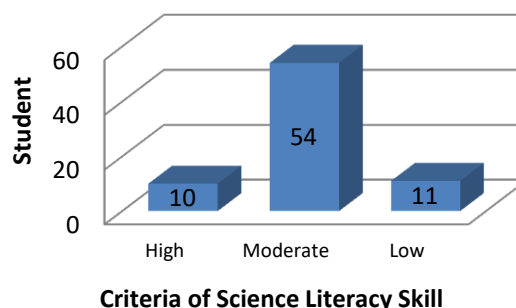


Figure 3. The Science Literacy Skill of Students

After getting an average score of scientific literacy, the scores are then grouped by the category of scientific literacy based on Table 7.

Table 7. Classification of Student Science Literacy

Criteria	Category
Nilai $>x + SD$	High
Nilai $>x - SD \leq \text{Nilai} \leq x + SD$	Moderate
Nilai $<x - SD$	Low

(Arikunto, 2011)

The results of the analysis of students' science literacy showed that students' science literacy skills are divided into three criteria, namely high, medium, and low. On average students' science literacy skills on the three different schools in middle category. Science literacy category was conveniently indicates the ability of students in Kudus district average middle category.

Science literacy skills students are assessed based on whether or not students choose the option of multiple choice questions and the reasons given. Each question has a score maximum number 2. The reason for choosing the answers were taken into account in this assessment because students must have consequences by giving arguments (Ridwan et al., 2013).

The analysis of the test instrument used by teachers there is a difference in scientific literacy based test instruments developed. The analysis was performed on an instrument with indicators of the same matter. Those indicators include the understanding of global warming and global warming. The test results achieved by students using the questions given by the teacher showed better results than the results of tests using instruments tesberbasis scientific literacy. This

difference is due to the problems that created the teachers have characteristics that directly refer to the issue in question, which were outlined clearly the data that will be used in the matter. This is different from the characteristics of matter based on scientific literacy developed. Problem-based scientific literacy has characteristics that requires students undertake an analysis of the issues presented through readings. Scientific data presented in the form of reading and is associated with an event or phenomenon that is close to student life. Problem-based scientific literacy also requires students to justify an answer, while the matter is made of teachers only multiple-choice and does not require the student to provide a reason. Different characteristics of these instruments to be one cause of the difference in the test results of students so that the students' science literacy measurement of average ability students in middle category. Scientific data presented in the form of reading and is associated with an event or phenomenon that is close to student life. Problem-based scientific literacy also requires students to justify an answer, while the matter is made of teachers only multiple-choice and does not require the student to provide a reason. Different characteristics of these instruments to be one cause of the difference in the test results of students so that the students' science literacy measurement of average ability students in middle category. Scientific data presented in the form of reading and is associated with an event or phenomenon that is close to student life. Problem-based scientific literacy also requires students to justify an answer, while the matter is made of teachers only multiple-choice and does not require the student to provide a reason. Different characteristics of these instruments to be one cause of the difference in the test results of students so that the students' science literacy measurement of average ability students in middle category.

Problem on aspects of the content have difficulty with the criteria level easy, medium and difficult. Indicator 1, indicator 3, and indicators 6 have the highest mean score between the mean scores on other indicators. Problems on the indicator 1, 3, and 6 had a moderate level of difficulty criteria. Problem with simple criteria implies that students have to understand the material in question and a lot of students who

answered correctly (Rahayu et al., 2014). Indicators 13 had the lowest mean score. Problem with difficult criteria can be caused by a statement or sentence in question is too long and complex so that students can not answer correctly because they do not understand the question (Djanuarsih, 2012). High difficulty level of questions that can not be used as a criterion that the matter is indeed difficult, but the question is actually easy but because students are less able to understand the concept then the matter becomes difficult to work (Rahayu et al., 2014). Problem with the criteria difficult to indicator 13 is a matter relating to the understanding of global warming.

Problem number 14 and 15 actually is no easy matter, but the average student less careful in understanding the graphs presented. Readings were quite long indeed a weakness in this instrument, because the problem with long sentences can make a student's difficulty in grasping its meaning (Djanuarsih, 2012). There are also students who can answer multiple choice right but can not give a reason. This could be because students guess the origin of the answer choices, but actually do not understand the concept so can not give a reason properly.

Problem on aspects of process/ competency explaining phenomena scientifically have a moderate level of difficulty with the criteria and difficult. Indicators 18 has a lower mean score because in Question 6 have a difficult criteria. Most students have difficulty in the process of identifying the molecules that cause global warming. These results are consistent with research Marsita et al. (2010) that one of the difficulties the students to the concept of differentiating molecules that cause global warming. Results of research conducted by Arief (2015) that the development of instruments to increase the competencies that exist in science literacy.

The students' science literacy skills as well as the study of the Holy PISA science literacy of students who said Indonesia is still low. Readings were quite long indeed a weakness in this instrument, karenasoal with a long reading can make students' difficulties in grasping the intent matter (Djanuarsih, 2012). There are also students who correctly answered multiple-choice correctly but can not give a good reason. This is because the students to guess the origin of the answer choices,

however, students do not understand the concept so can not give a reason properly.

Results obtained in line with OECD data (OECD, 2014) which put Indonesia ranked 64 out of 65 participating countries with an average acquisition value of science literacy component of Indonesian children by 382. In addition, Ridwan et al (2013) in his research the same conclusion, namely the ability of junior studied science literacy is at a functional level where the level is classified in science literacy skills are quite low.

The results of the analysis will serve as the scientific literacy ability determinant of the effectiveness of the assessment tool developed based on scientific literacy. Effectiveness based on scientific literacy assessment instrument based on a number of criteria for chemical literacy skills of students that can be brought out of the instrument. The results of students' science literacy skills analysis shows that there are criteria that varied capabilities. Each student is detected have different abilities. Therefore, based on scientific literacy assessment instrument is otherwise very effective because it can classify students' science literacy skills up to three criteria, namely low, medium, and high. The instrument also received a positive response from students and teachers.

CONCLUSION

Based on the results of the research and development, it can be concluded as follow: (a) the validity of the test instrument gets an average percentage of 79.99% and is included in the category of valid and feasible for use, (b) the practicality of test instruments that are in the category of very practical from the student responses 87.5 % and amounted to 94.23% of the teacher responses, (c) the effectiveness of the test instrument in the category effective, since it can reveal or develop the students' science literacy skills into three criteria, namely high, medium, and low. This study can be used as inputs and contributions to the efforts to improve the quality of manufacture of test instruments to measure the ability of scientific literacy on the theme of global warming.

The weakness of scientific literacy test instruments in this study is that it is not covering other aspects of the scientific literacy. Aspects that have not been included in this study could be

developed in future studies. In addition, the reading is too long on the instruments developed can be considered for further testing instrument development.

REFERENCES

- Akbar, S. (2013). *Instrumen Perangkat Pembelajaran*. Bandung: PT Remaja Rosdakarya.
- Arief, M. K., dan Utari, S. (2015). Implementation of Levels of Inquiry On Science Learning to Improve Junior High School Student's Scientific Literacy. *Jurnal Pendidikan Fisika Indonesia*, 11 (2), 117-125.
- Arikunto. 2011. *Dasar-dasar Evaluasi Pendidikan*. Jakarta: PT. Bumi. Aksara.
- Bashooir, K., & Supahar, S. (2019). Validitas dan reliabilitas instrumen asesmen kinerja literasi sains pelajaran fisika berbasis STEM. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 22(2), 219–230
- Dick and Carey. (1996). *The Systematic Design of Instruction*. Fourth Edition: Harper Collins College Publisher
- Djanuarsih, E. (2012). Validitas dan Reliabilitas Butir Soal. *E-Jurnal Dinas Pendidikan Kota Surabaya*; 1 (1), 1-12
- Fu'adah, H., Ani Rusilowati., Hartono. (2017). Pengembangan Alat Evaluasi Literasi Sains untuk Mengukur Kemampuan Literasi Sains Siswa Bertema Perpindahan Kalor dalam Kehidupan. *Jurnal Lembaran Ilmu Kependidikan*. 46 (1), 51-59
- Hobri. (2009). *Metode Penelitian Pengembangan (Developmental Research)*. <http://hobri.blog.ujec.co.id>
- Kurnia, F. Zulherman. & Fathurohman, A. (2014). Analisis Bahan Ajar Fisika SMA Kelas XI di Kecamatan Indralaya Utara Berdasarkan Kategori Literasi Sains. *Jurnal Inovasi dan Pembelajaran Fisika*, 1(1), 43-47
- Mardhiyyah, L., A., Ani Rusilowati, Suharto Linuwih. (2016). Pengembangan Instrumen Asesmen Literasi Sains Tema Energi. *Journal of Primary Education*. 5 (2), 147-154
- Marsita, R. A., S. Priatmoko, & E. Kusuma. (2010). Analisis Kesulitan Belajar Kimia Siswa SMA dalam Memahami Materi Larutan Penyangga dengan Menggunakan Two-Tier Multiple Choice Diagnostic Instrument. *Jurnal Inovasi Pendidikan Kimia*, 4(1), 512-520.
- Martin, M. O., Mullis, I. V. S., Foy, P., & Hooper, M. (2016). *TIMSS 2015 International Results in Science*. Diunduh dari Boston College, TIMSS & PIRLS International Study Center: <http://timssandpirls.bc.edu/timss2015/international-results/>

- Nisa, A., Sudarmin, dan Samini. (2015). Efektivitas Penggunaan Modul Terintegrasi Etnosains dalam Pembelajaran Berbasis Masalah untuk Meningkatkan Literasi Sains Siswa. *Unnes Science Education Journal*, 4(3), 1049-1056
- OECD. (2014). *PISA 2012 Results: What Students Know and Can Do (Student Performance in Mathematics, Reading and Sciences)*. Paris: OECD Publications.
- OECD.(2016). *PISA 2015 Results in Focus*. Paris: OECD.
- OECD.(2018). *PISA 2015 draft frameworks*. PISA, OECD Publishing, Paris
- Permanasari, A. (2011). Penerapan Project Based Learning Terintegrasi STEM Untuk Meningkatkan Literasi Sains Siswa Ditinjau dari Gender. *Jurnal Inovasi Pendidikan IPA*, 2(2), 202-212
- Rahayu, T. D., B. H. Purnomo, & Sukidin. (2014). Analisis Tingkat Kesukaran dan Daya Beda Pada Soal Ujian Tengah Semester Ganjil Bentuk Pilihan Ganda Mata Pelajaran Ekonomi Kelas X di SMA Negeri 5 Jember Tahun Ajaran 2012-2013. *Jurnal Edukasi UNEJ*, Vol 1, No 1, Hal. 39-43.
- Ridwan, M.S., L.A. Mardhiyyah & A. Rusilowati. (2013). *Pengembangan Instrumen dengan Pendekatan Kontekstual untuk Mengukur Kemampuan Literasi Sains Siswa*. *Seminar Nasional Evaluasi Pendidikan Tahun 2013*. Semarang: Universitas Negeri Semarang. Tersedia di <http://conf.unnes.ac.id/index.php/snep/I/paper/viewFile/23/17> [diakses 20-8-2019].
- Septiko, A., Wiyanto, S. E. Nugroho. (2015). Pengembangan Asesmen Untuk Mengukur Kemampuan Memecahkan Masalah, Bekerjasama Dan Berkomunikasi Calon Guru Fisika. *Journal of Innovative Science Education*, 4(2), 24-33
- Sudiatmika, A. I. (2010). Pengembangan Alat Ukur Tes Literasi Sains Siswa SMP dalam Konteks Budaya Bali. *Jurnal Penelitian dan Evaluasi Pendidikan*. 2(1), 1-40.
- Sugiyono.(2012). *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Sulistiawati. (2015). Analisa Pemahaman Literasi Sains Mahasiswa yang Mengambil Mata Kuliah IPA Terpadu Menggunakan Contoh Soal PISA 2009. *Jurnal Sainteks*, 12(1), 21-40.
- Sumarni, W., Sudarmin, Wiyanto, & Supartono. (2016). Preliminary Analysis of Assessment Instrument Design to Reveal Science Generic Skill and Chemistry Literacy. *International Journal of Evaluation and Research in Education (IJERE)*, 5(4),331-340