Asian Journal of Assessment in Teaching and Learning
Vol 13, Issue 2, 2023 (46-56) /eISSN ***2821-2916***

The Scientific Literacy Test Instrument on Particle Dynamics for High School Students

Mustika Wati[1*], Saiyidah Mahtari[1], Anis Muthi'ah[1], Dewi Dewantara[1], Suharno[2]
[1]Faculty of Teacher Training and Education, Universitas Lambung Mangkurat, Indonesia
[2]Faculty of Teacher Training and Education, Universitas Sebelas Maret, Indonesia

*Corresponding author: Mustika_pfis@ulm.ac.id

Received: 08 July 2023; **Revised:** 20 July 2023;
Accepted: 27 September 2023; **Published:** 21 December 2023

To link to this article: https://doi.org/10.37134/ajatel.vol13.2.5.2023

Abstract

The measurement of scientific literacy is very focused, this is shown by the urgency of preparing scientific literacy test instruments. This research develops a proper scientific literacy test instrument on particle dynamics material. This research is a development research adapted from the Borg and Gall model. The test subjects of this study were 202 students. Design validation, instrument validity, instrument reliability, instrument difficulty level and instrument differentiability are data analysis techniques. The product produced is a material science literacy test instrument for particle dynamics totaling 10 questions in the form of reasoned multiple choice questions. The results showed that (1) the questions were in the valid category, (2) had reliability with very good criteria of 0.83 (3) the difficulty level of the questions obtained 1 item in the difficult category and 11 questions with moderate criteria, (4) the results of the discriminating power of the questions show that there are 2 questions with sufficient discriminating power so that the items can still be used and there are 2 items with bias so that 2 items are discarded. It can be concluded that the developed scientific literacy test instrument can be used to measure students' scientific literacy competence.

Keywords: *Particle Dynamics, Test Instrument, Scientific Literacy, Validity, Reliability, Difficulty Levels, Instrument Differentiability*

INTRODUCTION

Technological developments make changes in the way of life and strong interactions in every human being. This encourages strength to be able to compete and adapt to become a quality resource. The educational aspect is closely related to quality human resources(Sholihah et al., 2023). One way to turn into a quality human being is to have life skills, one of which is scientific literacy (Al Sultan et al., 2018; Chusni et al., 2018; Dewantara et al., 2019; Tamassia et al., 2014; Winarni et al., 2020). Scientific literacy leads us to see events in nature and those resulting from human activities that are relevant to the concepts learned in school (Dewantara et al., 2020; Kholiq, 2020; Parno et al., 2020; Wahyu et al., 2020; Wartono et al., 2018; Yuriza et al., 2018; Zulistina & Sunarti, 2022). Students are expected to be able to explain scientific phenomena, evaluate and design scientific research, and interpret scientific data and evidence (Fakhriyah et al., 2017; Parno et al., 2020; Pentin et al., 2018). Scientific literacy is knowledge and understanding of scientific concepts and processes that enable a person to use knowledge to make decisions so that they can play an active role in all aspects of their lives (Cahyana et al., 2019; Hasasiyah, Siti Hardiyanti ; Hutomo, Bagus Addin ; Subali, Bambang ; Marwoto, 2019; Wahyu et al., 2020; Walag et al., 2022; Winarni et al., 2020)..

The importance of scientific literacy shows the urgency of having instruments to measure scientific literacy. The test or assessment instrument is a part of the planning and learning process. The assessment was carried out as an effort to measure the level of achievement of learning indicators and collect information on student learning developments in various aspects (Astuti et al., 2014). The test instrument is needed to determine the ability level of students and as a reference for developing policies about it (Milia et al., 2022). Several recent relevant studies have sought the preparation of scientific literacy instruments from various fields (Bashooir & Supahar, 2018; Pramuda et al., 2019; Suwono et al., 2022; Vizzotto & Mackedanz, 2018). In Indonesia, the development of scientific literacy instruments has also been implemented to test the scientific literacy of junior high school students in certain cities (Atta & Aras, 2020; Jufri et al., 2019; Septiani et al., 2019) and senior high school (Algiranto, 2022; Krisdiana et al., 2018; Lestari & Setyarsih, 2020; Novitasari & Handhika, 2018; Yamtinah et al., 2019; Yuliasih & Sarwi, 2020). Instruments have also been developed for prospective science teacher students (Rusilowati et al., 2019) as well as prospective physics teachers (Bahri et al., 2018). This is because in some cases, students in physics learning only focus on equations and mathematical problems without literacy. So, when learning begins to be intensified about literacy, physics teachers must also have instruments for assessing student literacy.

Instrument quality is focused on validity, reliability, level of difficulty, and distinguishing power. The validity of the test instrument is the accuracy of the test instrument in the form of description questions in measuring students' scientific literacy. Which is validated by the validator before being tested in the field. Instrument validity is an important part in a development. Validity determines the extent to which the instrument can measure the intended thing (Asyhari et al., 2019; Vizzotto & Mackedanz, 2018). The reliability of the test instrument is the extent to which the test instrument in the form of a description item is made credible whose results are not affected by circumstances. The reliability test was carried out to measure the level of consistency of the instrument being developed. A test is said to be reliable if it shows consistency from one test item to another (Novanti et al., 2018). Difficulty level is a measure of the ease of a question. Items with low criteria imply that students find it easy to solve problems, and items with difficult criteria imply that students find it difficult to solve problems (Adawiyah & Wisudawati, 2017). The comparison between easy, moderate, and difficult questions can be made 3-4-3, meaning that 30% of the questions fall into the easy category, 40% fall into the medium category, and 30% fall into the difficult category. Another comparison that is similar to the ratio above is 3-5-2. This means that 30% of the questions fall into the easy category, 50% fall into the moderate category, and 20% fall into the difficult category (Susanto, Hery ; Rinaldi, 2015). The differentiating power of the test instrument is the ability of the test instrument, in the form of descriptive questions, to distinguish high and low achieving students. Discriminating power is calculated using the number of right and wrong answers of students from the lower group and the upper group (Ismail, 2020). Discriminating power can be determined in two ways, namely differentiating into small groups (less than 100) and large groups (100 and above). In small groups, we divided the entire test group into two equal groups, 50% of the upper group and 50% of the lower group. For large groups, two poles are used to facilitate analysis. 27% each at the top and bottom only.

Table 1 Aspects of science literacy in the PISA 2015 assessment framework

Aspect	Indicator
Context	Current and historical personal, regional, national and global issues require the ability to understand science and technology.
Knowledge	Understanding various facts, concepts, and theories that underlie scientific knowledge including substantive, procedural, and epistemological knowledge.
Competence	The ability to provide explanations for scientific phenomena and interpret statistical data primarily based entirely on scientific evidence.
Attitude	A set of attitudes toward science characterized by an interest in science and technology, an appreciation of the scientific approach to inquiry, perceptions when present, and awareness of the environment.

Aspects of Scientific Literacy in the 2015 PISA assessment framework illustrated in Table 1. Therefore, in this study, researchers used the 2015 PISA aspects and indicators as a reference for developing instruments for testing basic material science literacy in particle dynamics for high school students. Based on the previous explanation, the purpose of this study is a description of research test instrument is needed that can be used to measure students' scientific literacy abilities.

METHODS

This type of R&D research is a process used to develop and validate educational products. The research and development model used by researchers is the Borg and Gall method (Borg & Gall, 1989), the steps can be seen in Figure 1.

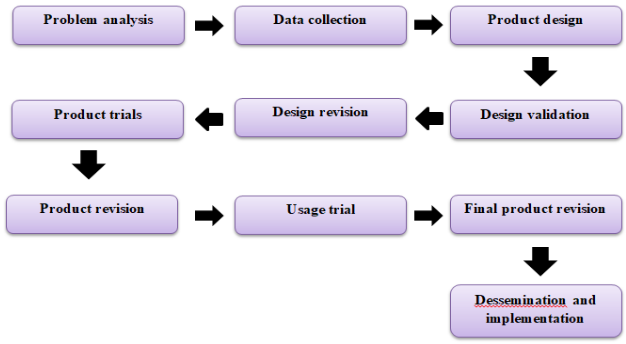

Figure 1 Steps for using the R & D method

The research took place from October 2021 to May 2022. The locations for the research were 3 school in Banjarmasin. The subject of this research is a material science literacy test instrument for particle dynamics. The test subjects were class X students for the 2021/2022 academic year, totaling 28 people for small product trials and 202 people for use trials. As for this study, several instruments were used, namely validation sheets and test instruments in the form of reasoned multiple choice which were given to students after one learning chapter had been taught. The development process carried out in this study only includes nine stages.

1. Problem analysis

Based on the results of the literature study, it was found that the scientific literacy abilities of Indonesian state students based on the results of the PISA assessment were still in the low category. The low ability of scientific literacy is due to several factors, one of which is that students are not used to working on scientific literacy-based questions and it is rare to find scientific literacy test instruments. Therefore, it needs an RND to develop assessment tools used by teachers when teaching physics in schools to find out the situation in the field.

2. Data collection

The results of the interviews were analyzed, then a literature survey was carried out as data collection to solve existing problems, and the tools developed met expectations. Data collection was carried out to collect references about scientific literacy and test instruments.

3. Product design

At this stage, it begins with determining the aspects and indicators of scientific literacy that will be used in research. Researchers compared aspects and indicators of scientific literacy used from several sources of literature. Then make a scientific literacy test instrument with 20 questions using the scientific literacy indicators. The scientific literacy test instrument that was prepared was reviewed by the supervisor. The design of the scientific literacy instrument which is compiled contains covers, prefaces, table of contents, aspects, and indicators of scientific literacy, question grids, and scientific literacy test questions. The scientific literacy test instrument was composed of 20 reasoned multiple-choice questions on class X high school particle dynamics material.

4. Design validation

The design validation stage was carried out by validators: 3 instructors (graduates) and 3 physics teachers (practitioners). This validation is based on rational thinking without field testing.

5. Design revision

The compiled products were repaired based on suggestions from the validator before being tested.

6. Product trials

The product trial phase was carried out in a small scope, namely on class X MIPA 3 students at SMA Negeri 6 Banjarmasin, totaling 28 people. The scientific literacy test instrument consists of 20 reasoned questions given to students. The trial lasted for 2 x 30 minutes or 2 JP.

7. Product revision

After trials were carried out on a small scale, the product was revised again to perfect the products made and take the questions received for trials on a large scale.

8. Usage trial

The usage trial phase was carried out on a large scale in three different schools. The scientific literacy test instrument consists of 12 reasoned questions given to students. The scientific literacy instrument was tested on 202 students consisting of 90 students from MAN 1 Banjarmasin, 99 students from SMA Negeri 6 Banjarmasin, and 13 students from SMA Islam Sabilal Muhtadin. The results of the data were then analyzed using the Rasch program in Winstep software version 3.73 to determine the value of validity, reliability, difficulty level, and item discriminating power.

9. Final product revision

Final product revisions were made to correct product deficiencies that remained after extensive application testing. This revision must improve the instruments developed according to conditions in the field. This phase provides the final output of the manuscript for the scientific literacy test instrument, a science test instrument consisting of 10 logical questions.

a. Data analysis technique

Design validity analysis was carried out using Aiken's V formula and then the value was adjusted according to the criteria $0.6 \leq V \leq 1.00$ in the valid category and $V < 0.6$ in the invalid category. The equations used are as follows:

$$V = \frac{\Sigma s}{[n(c-1)]} \quad (1)$$

$$s = r - l_0 \quad (2)$$

Meanwhile, the reliability value is calculated using the Alpha-Cronbach equation with the help of the SPSS application. Reliability assessment criteria can be seen in table 2. Interpretation of the reliability of the items. Calculation of the validity of reasoned multiple-choice items in small-scale trials using the product moment correlation formula, namely:

$$r_{XY} = \frac{N(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[N\sum X^2 - (\sum X)^2]\,[N\sum Y^2 - (\sum Y)^2]}} \quad (3)$$

Information:

r_{XY} : product-moment correlation coefficient
$\sum X$: the total score of all students on the item
$\sum Y$: the total score of all students on the test
N : the total number of all students
X : the score of each student on the item
Y : total score of each student

The criterion for testing the validity of the questions is after knowing the value r_{XY} and comparing it with *r* product moment with a table, if the price $r_{XY} > r_{tabel}$ then the question item is considered valid and otherwise $r_{XY} \leq r_{tabel}$ then the question is considered invalid. The reliability of the items is categorized in Table 2. Reliability can be calculated using the Alpha formula as follows:

$$r_{11} = \left(\frac{n}{n-1}\right)\left(1 - \frac{\sum \sigma_i^2}{\sigma_t^2}\right) \quad (4)$$

$$\sigma_i^2 = \frac{\sum x_i^2 \frac{(\sum x_i^2)}{N}}{N} \quad dan\ \sigma_t^2 = \frac{\sum x_t^2 \frac{(\sum x_t^2)}{N}}{N} \quad (5)$$

where

r_{11} : question reliability
n : number of items
$\sum \sigma_i^2$: the number of questions of the variance of the score of each item
σ_t^2 : variants total
$\sum x_i^2$: the total score of the question item number
$\sum x_t^2$: the total score of the questions
N : the number of students

Table 2 Criteria for the reliability of the items

Score	Criteria
$0{,}80 \leq r_{11} \leq 1{,}00$	Very high
$0{,}60 \leq r_{11} < 0{,}800$	High
$0{,}40 \leq r_{11} < 0{,}600$	Enough
$0{,}20 \leq r_{11} < 0{,}400$	Low
$0{,}00 \leq r_{11} < 0{,}200$	Very low

The difficulty level of objective questions can be determined using the following formula:

$$TK = \frac{\text{The number of students correctly answering the question item}}{\text{Number of students taking the test}}$$

with easy criteria (0.70 ≤ TK ≤ 1.00), medium (0.30 ≤ TK ≤ 0.70), and difficult (0.00 ≤ TK ≤ 0.30). The discriminating power of questions can be determined using the following formula:

$$DP = \frac{Upper\ group\ mean - Lower\ group\ mean}{maximum\ score\ of\ item}$$

Table 3 Criteria for discriminating power of questions (Wati & Miriam, 2017)

Discriminating Power	Criteria
$0{,}40 \le DP \le 1{,}00$	Accepted
$0{,}30 \le DP < 0{,}40$	Accepted, but needs improvement
$0{,}20 \le DP < 0{,}30$	Fixed
$0{,}00 \le DP < 0{,}20$	Not used/thrown away

Whether or not the analysis of test results using Rasch modeling is effective is the item obtained from the MNSQ score*(Mean Square Fit Statistics)* obtained by (0.5 < 0.5. MNSQ < 1.5), obtained the Z_{STD} *(Standardized Fit Statistic)* ($-2.0 < Z_{STD} < +2.0$) and the correlation value of the measured value (range 0.4 to 0.85) (Sumintono & Widhiarso, 2015). To measure the reliability of tests in the form of descriptive questions using the Rasch modeling, this is indicated by the individual separation values and items and scores *Alpha Cronbach.*

RESULT AND DISCUSSION

This instrument is assessed from several aspects such as material aspects, construction, language, and scientific literacy in which each aspect has 5 to 6 elaborations that are assessed from a score of 1 to 5. The validation of scientific literacy tools includes: (2) the field of composition, namely the clarity of the subject and scientific discourse; (3) language, including the proper and accurate use of Indonesian in interrogative sentences (Novanti et al., 2018). From the results of the data, it was found that the scientific literacy test instrument was included in the valid category. While the value of the reliability is high, which means that this instrument has a good level of trust. The reliability of the test instrument is the extent to which the test instrument is in the form of reasoned multiple-choice questions that are made credible and whose results are not affected by circumstances. Reliability is the accuracy or precision of a test being measured because the test is reliable if the results do not change over time (Wati & Miriam, 2017). An instrument is said to be valid if it can reveal data from variables accurately and does not deviate from the actual situation, while an instrument is said to be reliable if it can reveal reliable data (Nofriyandi ; Effenfi, 2019).

Then from the results of the design validation analysis, instrument improvements were made based on suggestions or input from the validator which included repairing the question sentences, writing procedures, and improvements regarding answer options. Overall this scientific literacy instrument is feasible to use after revision based on input from the validator. After the revision of the scientific literacy instruments, they were collected back into the validator and then approved for the next stage, namely product trials. The results of the validity and reliability of the two validators were as presented in Table 4.

Table 4 Results of the design validation analysis

No	Aspect	Average	Criteria
1	Theory	0.88	Valid
2	Construction	0.72	Valid
3	Language	0.85	Valid
4	Science literacy	0.75	Valid
Average		0.80	Valid
Reliability		0.96	Very high

Based the results of the Table 4, shows that the scientific literacy instrument has a valid category to be tested on students and its reliability is classified as high, so its reliability can be trusted to be assessed repeatedly to produce the same results. In addition, the results of this design validation are also improved according to input or suggestions from the validator. Then the small-scale product trial stage, the product trial aims to determine the validity, reliability, distinguishing power, and level of difficulty of the questions compiled. This product was tested on a small scale in class X senior high school who had studied particle dynamics with a total of around 28 students. The multiple-choice scientific literacy test instrument was used with a reasoned number of 20 questions. Test the validity of the items in this study using the product moment correlation formula or Pearson correlation. The number of students (N) = 28 and a significance level of 5%, the rtable = 0.374. The validity and reliability of product trials as presented in Table 5.

Table 5 Results of validity and reliability of product trials

No.	Question number	Criteria	
1	2, 3, 4, 5, 6, 7, 8, 10, 15	Height	
2	1, 9, 17	Enough	
3	18, 19	Low	
4	11, 12, 13, 14, 16, 20	Very low	
Valid	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 17	12 item	60 %
Invalid	11, 12, 13 14, 16, 18, 19, 20	8 item	40 %
Reliability	0,987924141		
Information	Very high		

Based on the results of the Table 5, it was found that of the 20 items compiled, there were 12 valid items and 8 invalid items. Items that are not valid because they include bad and very bad criteria and have a value of rxy < rtable or rxy < 0.374. Follow-up of 8 items whose validity is invalid then the items will be discarded or not reused. Whereas questions with valid validity can be reused during the product trial stage. A valid item means that it has an alignment between the question score and the total score, thus indicating that the item is capable of measuring students' abilities, whereas if the item score cannot provide support for the total score, the item is declared invalid (Indrawati & Sunarti, 2018). Test instruments that have invalid measuring power will not be able to provide information about students' abilities (Pratama, 2020). The items are said to be reliable because their reliability value in the criterion is very high, which is equal to 0.98. The difficulty level of the 20 questions tested on a small scale can be seen in Table 6. The results of Table 6 show that there are 12 items with difficult criteria, 6 questions with moderate criteria, and 2 questions with easy criteria. Based on these results, it can be seen that the scientific literacy test instrument is not proportional because the items with difficult questions are more difficult than the items with medium and easy criteria. The follow-up of the results of the difficulty level of the questions is items in the easy and difficult categories, re-examined to trace the causes of the items being difficult or easy to answer by students, after that they are corrected so that the questions can be used again in the next test. The results of the analysis of the differentiating power of the questions are shown in Table 7.

Table 6 The results of the analysis of the difficulty level of the questions

Question Number	Criteria	Percentage
7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19, 20	Difficult	60 %
1, 2, 4, 6, 8, 18	Currently	30 %
3, 5	Easy	10 %

Table 7 Results of the analysis of the differentiating power of the questions

Question Number	Criteria	Percentage
2, 3, 4, 5, 6, 8, 9, 10, 14, 15, 17, 18	Very good	60 %
1, 7, 19	Well	15 %
12, 13, 20	Enough	15 %
11, 16	Poor	10 %

Based on the results of the Table 7, it was found that the discriminating power of the 20 items resulted in four criteria, namely very good, good, fair, and bad. For questions that have very good discriminating power, good and sufficient, there are 18 items, and items that have bad discriminating power are 2 items. The follow-up carried out from the results of the differentiating power of the questions is that the items with poor discriminating power will be discarded. The questions were invalid because the answers from the students, the processing time was too short, namely 60 minutes for 20 reasonable multiple choice questions. The difficulty level of the test is not appropriate with the learning materials and teaching material resources available at school. The difficulty level of the invalid items is in the difficult category, while the discriminating power in items 11 and 16 is in a bad category, items 12, 13, and 20 are in the sufficient category, then items 14, 18, and 19 are in the very good category. Based on this, the invalidity of the items is influenced by the level of difficulty of the questions, then the invalidity of a coherent item with the distinguishing power of invalid items is on average in the bad and enough categories. Items 14, 18, and 19 will still have the distinguishing power included in the very good category of invalid item items this is due to the difficulty level of the questions and processing time. Follow-up of the effective results for this item includes reusing valid questions from a total of 12 questions in test use, canceling or removing 8 invalid questions, and reusing these questions in use trial use. The following is the distribution of questions based on the results of the analysis of the quality of the instrument.

The distribution of questions based on indicators consists of 5 items with indicators explaining scientific phenomena, 5 items with indicators of interpreting data and scientific evidence, and 2 questions evaluating and designing scientific investigations as shown in Table 8. As for the difficulty level, based on the previous analysis, there were 5 difficult items, 5 medium items, and 2 easy items out of 12 items. These items will be reused in the trial phase of use in three schools, namely MAN 1 Banjarmasin, SMA Negeri 6 Banjarmasin, and SMA Islam Sabilal Muhtadin Banjarmasin.

Table 8 Distribution of questions based on indicators of scientific literacy

Science Literacy Indicator	Question Number	Number of questions
Explaining Scientific Phenomena	1, 2, 5, 8, 9,	5
Interpret scientific data and evidence	3, 4, 6, 15, 17	5
Evaluating and designing scientific investigations	7, 10	2

Data from field trials were analyzed using Rasch modeling. Rasch modeling is considered valid or there are no elements extracted from the value (MNSQ) received <0.5. MNSQ < 1.5), the ZSTD value is obtained (*Standardized Fit Statistic*) (-2.0 < ZSTD < +2.0), and the correlation value of the measured values (range 0.4 to 0.85) (Chan et al., 2013; Sumintono & Widhiarso, 2015). Based on the results of data analysis, it was found that questions numbers 4 and 10 were biased, so they had to be aborted. Therefore, based on an analysis of the validity, reliability, difficulty level, and discriminating power of the questions to be aborted, there are 2 questions, namely questions numbers 4 and 10. The scientific literacy test instrument created has a reliability value of 0.83 which is a very good standard. The difficulty level of the scientific literacy test consists of 9 moderate questions and 1 difficult item. Thus, it can be concluded that the instrument developed is feasible to use, but the bias questions need to be removed so that the final stage of scientific literacy instruments will consist of 10 reasoned multiple choice questions that can be used as student learning tests. The existence of scientific literacy instruments can train and familiarize students with scientific literacy questions, so that if students are used to meeting scientific literacy-based test questions their scientific literacy skills will increase.

CONCLUSION

The product compiled is a scientific literacy test instrument on particle dynamics material which amounts to 10 reasoned multiple choice questions equipped with covers, prefaces, table of contents, aspects and indicators of scientific literacy, question grids, and scientific literacy questions. From validation sheets, product test results, and Rasch program results (MNSQ clothing scores, ZSTD clothing and measurement correlation values), the scientific literacy test instruments were validated. The scientific literacy test instrument created has a reliability value of 0.83 which is a very good standard. The difficulty level of the scientific literacy test consists of 9 moderate questions and 1 difficult item. The scientific literacy test instrument at the usage trial stage consists of 12 acceptable questions. However, there were two biased questions, so these two questions had to be discarded and the final result of the Science Literacy Tool was 10 questions. Thus, the questions that have been developed can be used by teachers as instruments for assessing scientific literacy in the classroom after learning, especially in senior high school particle dynamics material.

ACKNOWLEDGEMENTS

The author would like to thank the University of Lambung Mangkurat for assisting in the main research category through the "Dosen Wajib Meneliti" Program. The author also thanked the schools where the research was tested, namely SMAN 6 Banjarmasin, MAN 1 Banjarmasin, and SMA Islam Sabilal Muhtadin which provided the opportunity and trust to conduct research in schools. Not to forget, the author also thanked all class X MIPA students from the three schools who had participated in working on the test instrument so that the necessary data was obtained for this study.

FUNDING

The authors declare that no financial support was received for the research, authorship, and publication of this article.

DATA AVAILABILITY STATEMENT

Data will be made available on request.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Adawiyah, R., & Wisudawati, A. W. (2017). Indonesian Journal of Curriculum Pengembangan Instrumen Tes Berbasis Literasi Sains : Menilai Pemahaman Fenomena Ilmiah Mengenai Energi. *Indonesian Journal of Currculum and Educational Technology Studies*, *5*(2), 112–121.

Al Sultan, A., Henson Jr, H., & Fadde, P. J. (2018). Pre-Service Elementary Teachers' Scientific Literacy and Self-Efficacy in Teaching Science. IAFOR Journal of Education, 6(1), 25-41.

Algiranto, A. (2022). Pengembang Instrumen Tes Berbasis Literasi Sains Materi Fluida Statis Kelas XI SMA. *Silampari Jurnal Pendidikan Ilmu Fisika*, *4*(1), 47-58.

Astuti, W. P., Prasetyo, A. P. B., & Rahayu, E. S. (2014). Pengembangan Instrumen Asesmen Autentik Berbasis Literasi Sains Pada Materi Sistem Ekskresi. *Journal UNNES*, *43*(2), 94–102.

Asyhari, A., Fisika, P., Lampung, B., Lampung, B., Budaya, A., Asesmen, I., Sains, L., Agama, N., & Budaya, N. (2019). Pengembangan instrumen asesmen literasi sains berbasis nilai-nilai islam dan budaya indonesia dengan pendekatan kontekstual. *Lentera Pendidikan*, *22*(1), 166–179.

Atta, H. B., & Aras, I. (2020). Developing an instrument for students scientific literacy. *Journal of Physics: Conference Series*. https://doi.org/10.1088/1742-6596/1422/1/012019

Bahri, S., Simbolon, M., & Alhad, K. (2018). Development of Assessment Instrument of Scientific Literacy Ability for Students at Musamus University. *Proceedings of the International Joined Conference on Social Science (ICSS 2021)*. https://www.atlantis-press.com/article/125965194.pdf

Bashooir, K., & Supahar, S. (2018). Validitas dan reliabilitas instrumen asesmen kinerja literasi sains pelajaran fisika berbasis STEM. *Jurnal Penelitian dan Evaluasi Pendidikan*, *22*(2), 219–230. https://doi.org/10.21831/pep.v22i2.19590

Borg, W. R., & Gall, M. D. (1989). *Educational Research An introduction*. Longman.

Cahyana, U., Supatmi, S., Erdawati, & Rahmawati, Y. (2019). The influence of web-based learning and learning independence toward student's scientific literacy in chemistry course. *International Journal of Instruction*, *12*(4), 655–668. https://doi.org/10.29333/iji.2019.12442a

Chan, S. W., Ismail, Z., & Sumintono, B. (2013). A Rasch Model Analysis on Secondary Students' Statistical Reasoning Ability in Descriptive Statistics. *International Conference on Innovation, Management, and Technology Research*.

Chusni, M. M., Zakwandi, R., Hasanah, A., Malik, A., Ghazali, A. M., & Ubaidillah, M. (2018). Scientific Literacy: How is it Evolved to Pre-Service Physics Teacher? *Jurnal Ilmiah Pendidikan Fisika Al-Biruni*, *7*(2), 219–226. https://doi.org/10.24042/jipfalbiruni.v7i2.2781

Dewantara, D., Mahtari, S., & Haryandi, S. (2020). Validitas Lembar Kerja Mahasiswa Untuk Meningkatkan Literasi Sains Pada Mata Kuliah Fisika Biologi. *Kappa Journal*, *4*(1), 31-36.

Dewantara, D., Mahtari, S., Misbah, M., & Haryandi, S. (2019). Student Responses in Biology Physics Courses Use Worksheets Based on Scientific Literacy. *Prisma Sains : Jurnal Pengkajian Ilmu Dan Pembelajaran Matematika Dan IPA IKIP Mataram*, *7*(2), 192–197.

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

Hasasiyah, Siti Hardiyanti ; Hutomo, Bagus Addin ; Subali, Bambang ; Marwoto, P. (2019). Analisis Kemampuan Literasi Sains Siswa SMP pada Materi Sirkulasi Darah _ Hasasiyah _ Jurnal Penelitian Pendidikan IPA.pdf. *Jurnal Penelitian Pendidikan IPA*, *6*(1), 5–9.

Ismail, M. I. (2020). *Asasmen dan Evaluasi Pembelajaran* (Syariffudin, Ed.). Cendikia Publisher.

Jufri, A. W., Hakim, A., & Ramdani, A. (2019). Instrument development in measuring the scientific literacy integrated character level of junior high school students. *Journal of Physics: Conference Series*. https://doi.org/10.1088/1742-6596/1233/1/012100

Kholiq, A. (2020). Development of B D F-AR 2 (Physics Digital Book Based Augmented Reality) to train students' scientific literacy on Global Warming Material. *Berkala Ilmiah Pendidikan Fisika*, *8*(1), 50-58. https://doi.org/10.20527/bipf.v8i1.7881

Krisdiana, A., Aminah, N. S., & Nurosyid, F. (2018). The use of a four-tier wave diagnostic instrument to measure the scientific literacy among students in SMA negeri 2 Karanganyar. *Journal of Physics: Conference Series*. https://doi.org/10.1088/1742-6596/997/1/012042

Lestari, D., & Setyarsih, W. (2020). Kelayakan Instrumen Penilaian Formatif Berbasis Literasi Sains Peserta Didik Pada Materi Pemanasan Global. *IPF: Inovasi Pendidikan Fisika*, *9*(3), 561-570. https://ejournal.unesa.ac.id/index.php/inovasi-pendidikan-fisika/article/view/36287

Milia, A. H., Trisna, S., & Yanti, I. R. (2022). Development of HOTS Assessment Instruments on Static Electricity Materials for High School Level. *Berkala Ilmiah Pendidikan Fisika*, *10*(1), 73-83.

Nofriyandi ; Effenfi, L. A. (2019). View of Workshop Penyusunan Kisi-Kisi Soal Bagi Guru-Guru SMA PGRI Pekanbaru. *Community Education Engagement Journal*, *1*(1), 73–79.

Novitasari, L., & Handhika, J. (2018). Profil analisis kebutuhan pengembangan instrumen kognitif literasi sains untuk siswa SMA. *QUANTUM, Seminar Nasional Fisika Dan Pendidikan Fisika*. http://www.seminar.uad.ac.id/index.php/quantum/article/view/307

Parno, Yuliati, L., Hermanto, F. M., & Ali, M. (2020). A case study on comparison of high school students' scientific literacy competencies domain in physics with different methods: PBL-stem education, PBL, and conventional learning. *Jurnal Pendidikan IPA Indonesia*, *9*(2), 159–168.

Pentin, A., Kovaleva, G., Davidova, E., & Smirnova, E. (2018). Science education in Russia according to the results of the TIMSS and PISA international studies. *Voprosy Obrazovaniya / Educational Studies*, *2018*(1), 79–109. https://doi.org/10.17323/1814-9545-2018-1-79-109

Pramuda, A., Mundilarto, Kuswanto, H., & Hadiati, S. (2019). Effect of real-time physics organizer based smartphone and indigenous technology to students' scientific literacy viewed from gender differences. *International Journal of Instruction*, *12*(3), 253–270. https://doi.org/10.29333/iji.2019.12316a

Pratama, D. (2020). Analisis Kualitas Tes Buatan Guru Melalui Pendekatan Item Response Theory (IRT) Model Rasch. *Tarbawy : Jurnal Pendidikan Islam*, *7*(1), 61–70. https://doi.org/10.32923/tarbawy.v7i1.1187

Rusilowati, A., Yulianto, A., Ikhsannudin, Astuti, B., & Huda, N. (2019). Developing an instrument of scientific literacy assessment to measure natural science teacher candidates in force subject. *Journal of Physics: Conference Series*. https://doi.org/10.1088/1742-6596/1321/2/022027

Septiani, D., Widiyawati, Y., & Nurwahidah, I. (2019). Pengembangan Instrumen Tes Literasi Sains Berbasis Pisa Pada Aspek Menjelaskan Fenomena Ilmiah Untuk Siswa Kelas Vii. *Science Education and Application Journal*, *1*(2), 46-55.

Sholihah, A., Syahmani, S., & Suyidno, S. (2023). The Effectiveness of STEM Integrated Problem-Based Learning in Enhancing Student Science Literacy on Temperature and Heat Materials. *Jurnal Ilmiah Pendidikan Fisika*, *7*(1), 1–9. https://doi.org/10.20527/JIPF.V7I1.5639

Novanti, S. K. E., Yulianti, E., & Mustikasari, V. R. (2018). Pengembangan instrumen tes literasi sains siswa smp materi tekanan zat dan penerapannya dalam kehidupan sehari-hari. *Jurnal Pembelajaran Sains*, *2*(2009), 6–12.

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch pada Assesment Pendidikan*. Trim Komunikata.

Susanto, H., Rinaldi, A., & Novalia, N. (2015). Analisis Validitas Reliabilitas Tingkat Kesukaran Dan Daya Beda Pada Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Matematika Kelas XII Ips Di SMA Negeri 12 Bandar Lampung Tahun Ajaran 2014/2015. *Al-Jabar: Jurnal Pendidikan Matematika*, *6*(2), 203-218.

Suwono, H., Maulidia, L., Saefi, M., Kusairi, S., & Yuenyong, C. (2022). The development and validation of an instrument of prospective science teachers' perceptions of scientific literacy. *EURASIA Journal of Mathematics, Science and Technology Education*, *18*(1), em2068.

Tamassia, L., Frans, R., & Limburg, H. (2014). Does integrated science education improve scientific literacy? *Journal of the European Teacher Education Network*, *9*, 131–141.

Vizzotto, P. A., & Mackedanz, L. F. (2018). Validation of scientific literacy assessment instrument for high school graduates in traffic applied physics context. *Educação Em Revista*, *34, 1-36*. https://www.scielo.br/j/edur/a/7D5kV4RKHCx5gFjGStCSY5M/abstract/?lang=en

Wahyu, Y., Suastra, I. W., Sadia, I. W., & Suarni, N. K. (2020). The effectiveness of mobile augmented reality assisted STEM-based learning on scientific literacy and students' achievement. *International Journal of Instruction*, *13*(3), 343–356. https://doi.org/10.29333/iji.2020.13324a

Walag, A. M. P., Fajardo, M. T. M., Bacarrisas, P. G., & Guimary, F. M. (2022). A Canonical Correlation Analysis of Filipino Science Teachers' Scientific Literacy and Science Teaching Efficacy. *International Journal of Instruction*, *15*(3), 249–266. https://doi.org/10.29333/iji.2022.15314a

Wartono, W., Takaria, J., Batlolona, J. R., Grusche, S., Hudha, M. N., & Jayanti, Y. M. (2018). Inquiry-Discovery Empowering High Order Thinking Skills and Scientific Literacy on Substance Pressure Topic. *Jurnal Ilmiah Pendidikan Fisika Al-Biruni*, *7*(2), 139–151. https://doi.org/10.24042/jipfalbiruni.v7i2.2629

Wati, M., & Miriam, S. (2017). Pengembangan Instrumen Kognitif Literasi Sains Pada Pokok Bahasan Tekanan Di Kelas VIII SMP Kota Banjarmasin. *Jurnal Ilmiah Pendidikan Fisika*, *1*(3), 113–125.

Winarni, E. W., Hambali, D., & Purwandari, E. P. (2020). Analysis of language and scientific literacy skills for 4th grade elementary school students through discovery learning and ict media. *International Journal of Instruction*, *13*(2), 213–222. https://doi.org/10.29333/iji.2020.13215a

Yamtinah, S., Saputro, S., Mulyani, S., Ulfa, M., Lutviana, E., & Shidiq, S. (2019). Do students have enough scientific literacy? A computerized testlet instrument for measuring students' scientific literacy. *AIP Conference Proceedings*. https://doi.org/10.1063/1.5139875

Yuliasih, F., & Sarwi, S. (2020). Instrumen Penilaian Berbasis Keterampilan Abad Ke-21 untuk Mengukur Kemampuan Literasi Sains Siswa SMA. *UPEJ Unnes Physics Education Journal*, *9*(3), 320-330.

Yuriza, P. E., Adisyahputra, A., & Sigit, D. V. (2018). Correlation between higher-order thinking skills and level of intelligence with scientific literacy on junior high school students. *Biosfer*, *11*(1), 13–21.

Zulistina, H., & Sunarti, T. (2022). Analyzing Islamic Boarding School Students' Scientific Literacy Skills on Physics Material. *Prisma Sains: Jurnal Pengkajian Ilmu Dan Pembelajaran Matematika Dan IPA IKIP Mataram*, *10*(2), 361–372.