



# Online-based Literacy Test Instruments for Evaluating Prospective Elementary School Teachers' Science Competencies

Tri Wiyoko<sup>1\*</sup>, Titis Wulandari<sup>2</sup>, Muhammad Irhas<sup>3</sup> 

<sup>1,3</sup> PGSD, Muhammadiyah University Muara Bungo, Bungo, Jambi, Indonesia

<sup>2</sup> PJKR, Muhammadiyah University Muara Bungo, Bungo, Jambi, Indonesia

## ARTICLE INFO

### Article history:

Received September 20, 2023

Accepted April 06, 2024

Available online May 25, 2024

### Kata Kunci:

Literasi Berbasis Online,  
Instrumen Tes, Kompetensi IPA

### Keywords:

Online-based Literacy, Test  
Instrument, Science  
Competencies



This is an open access article under the  
CC BY-SA license.

Copyright © 2024 by Author. Published  
by Universitas Pendidikan Ganesha.

## ABSTRAK

Asesmen Kompetensi Minimal digunakan untuk mengukur kinerja sekolah berdasarkan literasi dan numerasi. Berkaitan dengan hal tersebut, calon pendidik siswa sekolah dasar perlu melatih dan menguasai kompetensi literasi sains. Penelitian ini bertujuan untuk menghasilkan instrumen tes literasi berbasis online untuk menilai kompetensi sains siswa pendidikan guru sekolah dasar. Jenis penelitian ini adalah penelitian dan pengembangan. Prosedur pengembangan instrumen tes literasi menggunakan model 4D (Define, Design, Development dan Discussion). Penelitian ini melibatkan 3 validator dan subjek uji empiris yang terdiri dari 30 mahasiswa PGSD. Pengumpulan data menggunakan lembar analisis kebutuhan, lembar validasi bahasa, materi dan grafik. Selanjutnya dilakukan validitas empiris dengan menggunakan program Quest untuk mengetahui tingkat kesulitan (Thresholds) dan kesesuaian butir soal (Outfit-t). Hasil validasi instrumen soal tes menunjukkan bahwa 89,5% validator materi berada pada kategori sangat baik, 92% validator bahasa berada pada kategori sangat baik, 96,5% validator grafis berada pada kategori sangat baik. Sebaran soal yang dikembangkan menunjukkan kemampuan siswa dalam mengidentifikasi permasalahan ilmiah sebesar 66% berada pada kategori sedang, 65% dalam menjelaskan fenomena ilmiah berada pada kategori sedang dan kemampuan menggunakan bukti ilmiah sebesar 45% berada pada kategori sangat rendah. kategori. Jadi instrumen soal tes yang dikembangkan memenuhi validitas materi, bahasa, kegrafikaan dan validitas empirik, sehingga layak digunakan untuk mengukur kompetensi literasi sains siswa.

## ABSTRACT

The Minimum Competency Assessment is used to measure school performance based on literacy and numeracy. In this regard, it is necessary for prospective elementary school student educators to train and master scientific literacy competencies. This research aims to produce an online-based literacy test instrument for assessing the science competency of elementary school teacher education students. This type of research is research and development. The procedure for developing literacy test instruments uses the 4D model (Define, Design, Development and Discussion). This research involved 3 validators and empirical test subjects consisting of 30 Prospective Elementary School Teachers. Data collection uses needs analysis sheets, language validation sheets, materials and graphics. Next, empirical validity is carried out using the Quest program to determine the level of difficulty (Thresholds) and suitability of the items (Outfit-t). The results of the test question instrument validation show that 89.5% of the material validation is in the very good category, 92% of the language validator is in the very good category, 96.5% of the graphic validator is in the very good category. The distribution of questions that have been developed shows that students' ability to identify scientific issues is 66% in the medium category, 65% in explaining scientific phenomena is in the medium category and the ability to use scientific evidence is 45% in the very low category. So the test question instrument developed meets the validity of the material, language, graphics and empirical validity, so it is suitable to be used to measure students' scientific literacy competency.

## 1. INTRODUCTION

The development of education moves very quickly and complexly. The carrying capacity of existing technology has transformed education towards openness in learning. This is due to the magnitude of the challenge to meet the quality needs of human life. Therefore, it is very important to incorporate 21st century skills into science education to form knowledge and understanding of scientific concepts and processes necessary for personal decision making and social life (Abusa'aleek & Baniabdelrahman, 2020; Turiman et al., 2012). One of the skills needed in the 21st century is scientific literacy, because scientific literacy views the importance of thinking and acting skills which involve mastering thinking and using scientific thinking in recognizing and responding to social issues (Pratiwi et al., 2019; Singh, 2019). Education in the 21st century promotes learning with a mind and problem-solving orientation, independence in learning leads to character honing and qualified social life in the context of tolerance and cooperation (Izhar et al., 2022; Singh, 2019). The independent learning policy has determined that literacy is one of the assessment aspects applied in schools. Scientific Literacy is the ability to identify, understand, and interpret issues related to science that are needed by someone to make decisions based on scientific evidence (Prasetya & Adlan, 2022; Sulasmi, 2022). Science literacy skills must be mastered by prospective elementary school teachers. This is due to the importance of prospective teachers playing a role in the world of education with the obligation to be information literate and to respond quickly to changes in educational policies. Integrated scientific literacy implementation will color students' scientific experiences and opportunities to understand, understand, and make sense of the relationship between science, technology, and society which in turn will affect their personal life, career, and future. Science literacy for students can be developed through exercises or habituation in the learning process provided by the teacher's teaching. So in the 21st century, prospective elementary school teachers are needed who can master scientific literacy skills.

Study program students education school base as prospective school teacher base already should have Skills good scientific literacy, because ability scientific literacy is used as stock for transfer skills and knowledge related with science. Furthermore, scientific literacy is closely related to students' ability to understand the environment, life, and problems in modern society which is very dependent on technology and the development of science or knowledge (Bybee & McCrae, 2011; Siahaan et al., 2022). Strong scientific literacy is related to the ability of students of to understand the environment, life, and problems in a highly dependent modern society with technology as well as the development of knowledge or science. because a teacher should already equip himself to have adequate competence in Duty guiding, coaching, and directing participant teachers to have mastery of scientific literacy. Aspect's own scientific competence refers to the mental processes involved when answering something question or solving problem (Toharudin, 2011; Widowati et al., 2019).

Priority PISA assessment 2012 in scientific literacy focuses on several aspects of scientific competence, namely: identifying issues scientifically, explaining phenomena scientific based on scientific knowledge, and using scientific proof (Clarke, 2020; Özer & Bilgisi Öz, 2020). Assessment results the 2018 Program For International Student Assessment (PISA) shows that participant's scientific abilities students in Indonesia on average reach a score of 389 with the OECD average score is 489. These results classify Indonesia in the quadrant of low performance with high equity. Next, there is a gender gap in performance inequality performance studies Among girls and boys. Participants educate women more well than participants educate men in all fields in PISA (Bybee & McCrae, 2011; Novita & Putra, 2016). In general, Indonesian students are not yet accustomed to scientific literacy learning, this is because science learning is not yet oriented towards developing scientific literacy as a competency that must be possessed. So students are generally less trained in solving problems such as indicators in scientific literacy (Rubini et al., 2018; Widowati et al., 2019). As for the steps that can conducted to repair the existing situation, that is with applying Assessment Minimum Competency (AKM).

AKM implementation is used to measure performance in school based on literacy and numeracy. About Thing it, the need for student candidate educators in elementary school to train and master ability literacy is one of them, namely competence scientific literacy. Science learning is expected to develop students' scientific literacy so that it can improve the quality of science learning (Jufrida et al., 2019; Kelana et al., 2021). Science learning is expected to develop students' scientific literacy so that it can improve the quality of science learning (Suryawati et al., 2018; Udompong et al., 2014). Scientific literacy is scientific knowledge and skills to be able to identify questions, obtain new knowledge, explain scientific phenomena, and draw conclusions based on facts, understand the characteristics of science, awareness of how science and technology shape the natural, intellectual and cultural environment, and the willingness to be involved and care on science-related issues.

Ability scientific literacy can be measured by the instrument questions that are designed based on existing indicators. Based on the reasons that have been described, there is a need to develop an instrument to test literacy online based on the evaluation of the science competency of Prospective Elementary School

Teachers. The aim of this study is to produce an online-based literacy test instrument for assessing the science competency of elementary school teacher education students. The novelty of this study is necessary for prospective elementary school student educators to train and master scientific literacy competencies.

## 2. METHOD

Research that has been held, including in-type research and development (Research and Development). Developed product form Instrument question test online based on the use of Google Forms app with destination for used in evaluating science competency of Prospective Elementary School Teachers. Procedure development research used is a 4D model (Define, Design, Development, and Dissemination ), a development model developed by (Richey & Klein, 2014; Thiagarajan, 1976).



Figure 1. 4D Development Design

At the stage beginning of development has conducted activity definition that includes aspect analysis needs analysis of competence base for determining achievements competence learning and determining indicator scientific literacy for producing instrument question test scientific literacy. Stage planning is conducted with planning instrument question test form grid matter to be made as guidelines in developing instrument question test literacy online so you can measure science competency. Stage development is conducted by making reading as scientific literacy and question-related tests with appropriate topics with science. The questions that have been asked are validated by experts. The results of the assessment by experts used as a base is instrument question developed test worthy or not. At this stage, this question test can be revised and returned in accordance with suggestions from the validator. Stage deployment instrument question test literacy online will be implemented on a wide scale for Prospective Elementary School Teachers. The results of the work will used to analyze the evaluation of competence science from Prospective Elementary School Teachers.

Data collection using sheet observation analysis requirements, validation sheet Instrument questions and questions Test Scientific literacy in the form of choice served double online. The data analysis technique used in a study is analysis qualitative which gives an analysis of the data obtained and describes the data that is qualitative. Data Validation Test results instrument question tests that have been developed based on criteria as used by. The percentage score obtained is then transformed in Table 1.

Table 1. Ranges Percentage and Criteria Evaluation Validators

Intervals	Criteria
81 % < score ≤ 100%	Very Well
61% < score ≤ 80%	Well
41% < score ≤ 60%	Enough Well
21% < score ≤ 40%	Not enough Well
0% < score ≤ 20%	Very No Well

Study this set appropriateness instrument question test with criteria good as results evaluation from the Material /Content validator, the Eligibility Validator language and presentation/graphic validators. If the results evaluation ends with a minimum value that is good, then the developed product is worthy of use. Next, test results \_grain question test used for knowing the validity empirical from development instrument test questions. The results of the answers obtained were analyzed with the use of theory modern tests with a quests program for knowing level difficulty (Thresholds) and suitability grain questions (Outfit-t).

Difficulty-level grain questions show the magnitude possibility of how many respondents can answer a grain question with the right (Setyawarno, 2017). The level of difficulty of the items through the QUEST program is seen from the value (threshold) item estimate with criteria as show in Table 2.

**Table 2. Provisions Score Threshold**

Threshold Value	Information
$b > 2$	Very Difficult
$1 < b \leq 2$	Hard
$-1 < b \leq 1$	Currently
$-1 > b \geq -2$	Easy
$b < -2$	Very Easy

**Compatibility Grain Question (Out Fit -t)**

Compatibility item (item fit) is information for determining if grain is an appropriate item or not by measuring the ability of students for. grain questions that don't fit eliminated so as not to influence quality evaluation. As for the items that fit can used and developed for evaluation more carry-on (Setyawarno, 2017; Suparman, 2020). The of fit each item in the QUEST program is determined based on the magnitude MNSQ or INFIT value, the INFIT value of the item in question. Determination of Item Fit in each item question is show in Table 3.

**Table 3. MNSQ INFIT Provisions for the Rash Model**

MNSQ INFIT Score	Information
$>1.33$	Not suitable with the Rasch model
$0.77 - 1.33$	Suitable with the Rasch model
$< 0.77$	Not suitable with the Rasch model

After the question is validated in a manner of content and empirical validation, the question is spread for testing (Rahmadani et al., 2018). The result of data processing student from grain questions that have been developed will analyzed and interpreted based on Table 4.

**Table 4. Categories Competence Student Scientific Literacy**

Intervals (%)	Category
86-100	Very High
76-85	High
60-75	Currently
55-59	Low
$\leq 54$	Very Low

**3. RESULT AND DISCUSSION**

**Result**

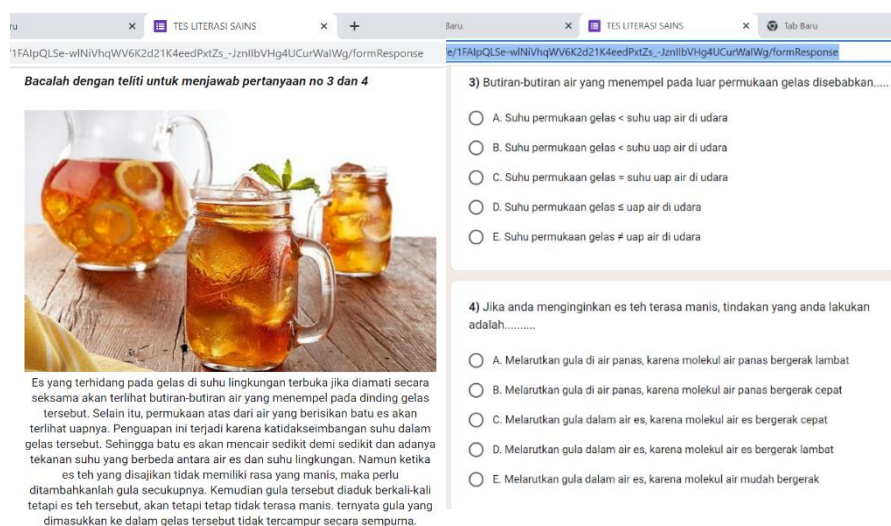
Study development that has implemented in develop instrument test literacy online based for evaluating student science competence obtained results and discussion. At the definition stage, a needs analysis is carried out through interviews with lecturers. Interview results inform that the questions tested on students so far this not yet deep for reveal ability scientific literacy. The question given during this test is completeness understanding student reviewed from aspect understanding, yet reach the analysis and application of the components in life every day. Next, the questions tested are also in deep offline form. There is a development question online-based scientific literacy, of course, will give innovation in evaluation learning draft basic science. After analysis needs to done, next that is analyses of competence determine achievements learning from questions to be developed. Analysis of learning outcomes based on the semester learning plan that has been made with the lecturer. As for achievements learning to be developed that is from Theory changer exists objects, forces and motion. Next achievements selected learning harmonized with indicators from scientific literacy.

At design stage instrument question test for evaluation competence developed scientific literacy in form question choice double as many as 15 items. Developed questions started with discourse or related reading with Theory learning. The grid results development is presented in Table 5.

**Table 5. Question Grid Test Science Literacy**

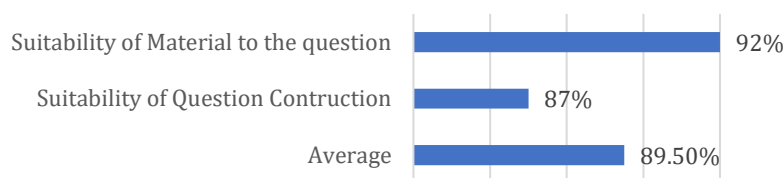
Theory	Achieved Learning	Indicator Science Literacy	Number Question
Change Shape of Things	Students capable of differentiating change material and changing an existing object	Identify issue scientific	1, 2, 3
	Student capable of explaining reason change Physics	Explain phenomenon scientific	4, 5,6,7,8
Style and motion	Student capable explain reason for changing the chemical	Use proof scientific	9, 10,11,12 13,14, 15
	Student capable differentiate substance single and mixed		
	Able to analyze draft style and influence subject matter style		
	Students capable of analyzing resultant style and influence on subject matter style		
	Students capable of identifying the application of Newton's law in life daily		

Next, questions that have been made, then moved to Google form so you can access them by students online. The question can be accessed at the following link <https://forms.gle/X3bTbNaBtkuTNR4u9>. The next appearance question of scientific literacy is presented on Google form as show in Figure 2.



**Figure 2. Display Question in Google Forms**

At stage development, conducted assessment by experts who are competent in their fields i.e., field validator IPA material, language validator, and graphic validator. This is for evaluate appropriateness from the question instrument choice developed double. The material validator consists of two aspects: material and construction. Aspect Theory consists of 5 indicators and aspects construction consists of 9 indicators. As for detailed results, the material validator assessment is show in Figure 3.



**Figure 3. Material Expert Assessment**

From Figure 3, whole indicators assessed on the aspect suitable material with question obtained percentage by 92% with very good criteria. Then, for suitability, construction questions with questions obtained a percentage of 87% with very good criteria. Average yield evaluation expert Theory by 89.5%. this shows that questions of scientific literacy created worthy for test student science competence. Next for

evaluation from the Language aspect includes 5 indicators that is suitability with development students, Communicative, Dialogic and Interactive, Coherence, and coherence think, fit with Indonesian rules with right. Based on 5 indicators mentioned, obtained the assessment of the Indonesian language expert validator is presented in Figure 4.

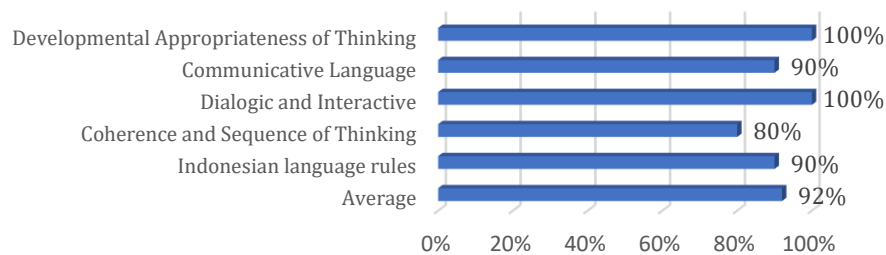


Figure 4. Assessment Language Aspect

Based on Figure 4 of the Indonesian validator assessment, quality developed questions with criteria appropriate question with development method think student by 100% with very good category. Next, because it's also communicative, with a percentage of 90% in the very good category, dialogue and interactive 100% in the very good category. Use coherence and coherence question practice method think student 80% in very good category. Final suitability with Indonesian rules with Correct by 90% with very good criteria. The average validator rating of the language aspect is 92%, with a very good category. So, matter developed scientific literacy from aspects of Indonesian have quality worthy for test student science competence. Evaluation graphics from grain-developed questions need validation. this is because there is several image used to clarify related phenomena with grain question. Assessment results from the validator are presented in Figure 5.

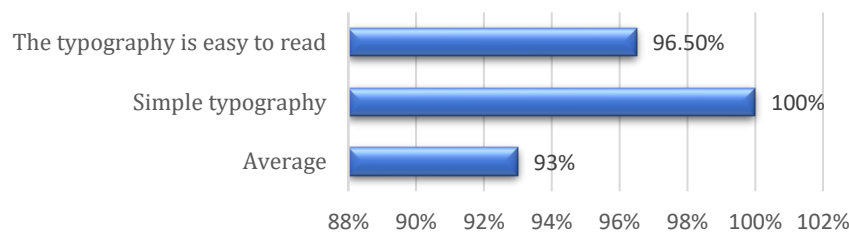


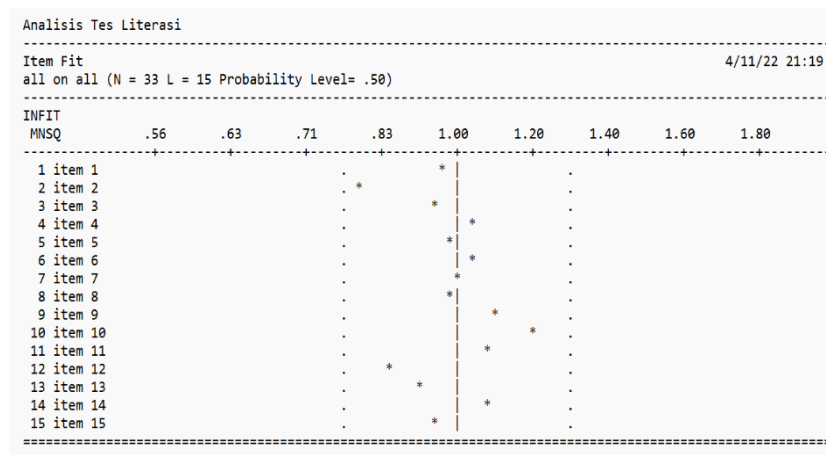
Figure 5. Graphic Expert Assessment

Based on Figure 5 obtained the results evaluation graphics with percentage typography simple by 93% which is a very good category. Then, for easy typography, I obtained a percentage of 100% with very good categories. Average yield evaluation graphics by 96.5% with very good category. So, Grain developed questions for aspect graphics worthy for tested try right. Grain already matters validated by the material validator, language, and graphics, then next with validity empirical validity using the Quest program. Item test analysis questions using the Quest program on Thresholds show existing index hardship grain questions. Obtained data analyzed with theory modern tests use logistics one parameter (1PL) because the Quest program only can analyze one parameter logistic model. The test results using Quest obtained data in Table 6.

Table 6. Summary Category Threshold (b) Analysis Test Literacy

Threshold value	Information	Number Grain	Total
$b > 2$	Very Difficult	14	1 (6.67%)
$1 < b \leq 2$	Hard	3,5,9,15	4 (26.7%)
$-1 < b \leq 1$	Currently	6,7,8,10,11,12,13	7 (46.7%)
$-1 > b \geq -2$	Easy	1,2	2 (13.3%)
$b < -2$	Very Easy	4	1 (6.67%)

Grain already matter analyzed Thresholds, continued with analyzing output from the Quest program on INFIT MNSQ. Output data from the program Quests declared fit if the item or test case person complied coefficient range limit INFIT MNSQ from 0.77 - 1.30. As for output results from the grain, the question test developed literacy after testing it served on [Figure 6](#).



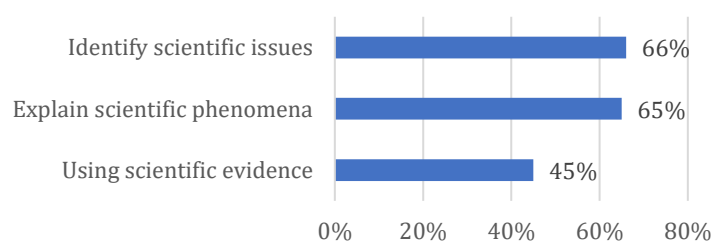
**Figure 6.** Fit Map items for Analysis Test Literacy

Based on [Figure 6](#), the INFIT MNSQ output from whole developed questions there are 15 items question test adequate literacy coefficient in fit, meaning 100% grain question test literacy his Fulfill coefficient between 0.77- 1.30. Grain questions that are in fit above 88.5%, then also did prolific research percentage compatibility grain test with each model by 88%. Grain percentage above 80 % indicates that a large evaluation item is suitable for use for evaluation learning. Outfit items result is show in [Table 7](#).

**Table 7.** Outfit Items

Intervals	Category	Item Number	Average Outfit-T	Information
Outfit $t \leq 2.00$	Pass Question	1,2,3,4,5,6,7,8,9, 10,11,12,13,14,15	1.04	Suitable with the Rasch model
Outfit $t > 2.00$	Question Fall			

The data shown in [Table 7](#), shows that the 15-item question tests appropriate scientific literacy with the Rasch model, the Outfit-t average is 1.04. So the problem is developed Fulfill validity material, language, graphics, and validity empirical. Deployment product form question test already validated in a manner content and empirical, done with give question online for test ability scientific literacy of semester I students in academic 2022/2023. The results obtained are presented in [Figure 7](#).



**Figure 7.** Test Results Student Scientific Literacy

Test results literacy from grain developed questions based on [Figure 7](#), shows that ability literacy student varies. Student science competency for Indicator identifies issues scientific by 66% with category medium, on the indicator this ability low student that for knowing characteristic features key from investigation scientific. Then for indicator explain the phenomenon scientifically by 65% with category medium. Ability students who are low on the indicator, that is, describe or interpret phenomenon and predict change. Next, for the ability to use scientific proof by 45% with a very low category. The ability of students to use proof science needs to be improved in the aspect of interpreting proof scientific and

interesting conclusions and giving reflections based on the conclusion. Based on the score percentage above, there is 1 indicator with a very low category, while the other 2 indicators have reached category medium.

## Discussion

Scientific literacy is scientific knowledge and skills to be able to identify questions, obtain new knowledge, explain scientific phenomena, and draw conclusions based on facts, understand the characteristics of science, awareness of how science and technology shape the natural, intellectual and cultural environment, and the willingness to be involved and care on science-related issues (Holbrook & Rannikmae, 2009; Lederman et al., 2013). Evaluation of scientific literacy can be done by testing individual knowledge through literacy question instruments. The development of scientific literacy instrument questions begins with the process of defining, design, development and dissemination. The development process is carried out by paying attention to the question grid which includes question criteria that represent the content of the curriculum/competencies. The components are detailed, clear and easy to understand and the questions can be made in accordance with the specified indicators and question form (Firmansyah & Gradini, 2020; Suparman, 2020). The literacy question instrument developed was applied to the Googleform website. selecting web media in the form of a Google form will make the use process easier. Apart from that, it shows that web media has a positive effect on scientific literacy, students have higher learning independence compared to traditional media (Cahyana et al., 2019; Elmahdi et al., 2018).

Furthermore, the development of literacy question instruments also considers material selection, because mastery of science material can contribute to the ability to solve questions. This shows that there is a correlation between the level of relationship between concept mastery and problem solving abilities for prospective teacher students (Batlolona et al., 2018; Hakim et al., 2022). The validation results from the material aspect show that very good criteria were obtained with a percentage of 89.5%, making it feasible to apply. Furthermore, the use of good and correct language in questions is very helpful in solving problems. The use of language is not only for interacting with each other, but good command of language in questions in accordance with existing concepts and procedures makes it easier to find solutions to problems (Abdullina et al., 2019; Walkington et al., 2019). The validation results show that the language used in developing the literacy test questions is at very good criteria or 92%, so the literacy questions developed are suitable for use. Apart from material and language, typographic aspects need to be validated. Through the use of text in literacy matters, typography plays an important role because it conveys the essence of effective communication. Apart from that, typography in visual media can create interactive and interesting material and build communication between teachers and students (Sulaiman et al., 2020; Wasito et al., 2022). Thus, a good literacy question instrument must fulfill the construct validity of the question.

The validity of developing literacy test questions is also reviewed from the difficulty index of the questions. The characteristics of questions with a good level of difficulty have a composition of easy, medium and difficult questions distributed proportionally according to the material being tested and a good item differentiation index is able to differentiate between high ability groups and low ability groups (Kibirige & Teffo, 2014; Sarea & Ruslan, 2019). The index hardship grain very diverse questions. grain matter that has very difficult criteria ( $b > 2$ ) of 6.67% which is on question number 14. Item question with criteria hard  $1 < b \leq 2$ ) of 26.7% or 4 points question. next grain question with criteria currently ( $-1 < b \leq 1$ ) of 46.7% or 7 points questions and criteria easy ( $-1 > b \geq 2$ ) as much as 13.3% or 2 points question. whereas grain questions with very easy criteria ( $b < -2$ ) as much as 1 point question or 6.67%. The characteristics of the evaluation instrument developed are in line with previous research that the instrument has a level of difficulty analysis shows that almost all questions have good parameters ( $b_1 < b_2 < b_3$ ), with a difficulty index ranging from -0.332 to 0.144 and categorized as moderate (Martinah et al., 2022; Saepuzaman et al., 2022). Furthermore The validity of the questions obtained results of 70%, which means it is said to be valid, There are 30% of questions in the medium category (not difficult and not easy) and questions with good distinguishing power.

Test question could be a good test if it has a level of no difficulty, too hard but also not too easy (moderate) (Friatma & Anhar, 2019; Ramadhan et al., 2023). Whereas grain easy matter must repaired moreover, item question with category easy he delivered question number 1 is used for test related scientific competencies with know possible issues investigated in a manner scientific. Next grain question number 2 for test competency identifies keywords for information scientific. Grain question number 4 is used to test deep scientific competence and apply deep scientific knowledge in a given situation. Items must repaired no matter how difficult it is no, will motivate participants to be educated to look for effort in finding solutions to the problem. Whereas matter that has a level of difficulty will make participants break hope and no spirit in looking for settlement problem at hand (Saepuzaman et al., 2022; Suharsimi Arikunto, 2021).



Dissemination results grain question existing literacy implemented give description ability student science competence. Student science competency for prospective school teachers base is very important for improvement because good scientific competency will run a good learning process inside the class. Therefore, scientific literacy is very important for solving various related problems with ethical, moral, and global issues of consequence rapid change in science and technology (Adhani et al., 2020; Yuliati & Saputra, 2019). As for the efforts that can conducted for enhancement student science competency that is with arrangement room scope given material and process activities learning done. Apart from that, measuring scientific literacy is also important, because determining students' literacy in scientific concepts will train students' level of thinking, as well as understanding scientific processes and applying them to problems (N.M.S. Yantiningsih et al., 2022; Saputra & Sujarwanta, 2021). Furthermore, the use of scientific literacy tests can encourage alignment of learning objectives, class activities, and student assessments in uncovering challenges and alternative conceptions in using scientific literacy skills (Cole & Feng, 2015; Gormally et al., 2012). With this ability, student science competency needs to be measured and corrected for it to science.

Thus, the development of scientific literacy questions can be used as an alternative evaluation instrument in collecting data and a tool for obtaining information about students' scientific literacy abilities. Furthermore, students can be trained to improve their scientific literacy skills. So that the quality of students' scientific literacy can be better in terms of their ability to identify scientific issues, explain scientific phenomena and use scientific evidence.

#### 4. CONCLUSION

The results of the validation of the online-based literacy test instrument developed using the 4D model concluded that there were 15 literacy test questions that were suitable for use to assess the science competency of elementary school teacher education students. The validation results of the test question instrument show that the validation scores for material, language and graphics are in the very good category. Furthermore, the empirical validation results show that the 15 questions developed meet the INFIT MNSQ coefficient range limit, with an average Outfit-t of 1.04 so that it is in accordance with the Rasch model. So the test question instrument developed meets the validity of the material, language, graphics and empirical validity, so it is suitable to be used to measure students' scientific literacy competency.

#### 5. ACKNOWLEDGE

Thank you to LPPM Muara Bungo Muhammadiyah University for providing internal research funding grants. Funding for this internal research is based on contract number 0924.07/14/LP3M.STKIP-MB/II/2022.

#### 6. REFERENCES

- Abdullina, L. R., Ageeva, A. V., & Gabdreeva, N. V. (2019). Using The Flipped Classroom Model in The Teaching of The Theoretical Disciplines (French Language) at The University [Technologie De La Classe Inversee Dans L'enseignement Des Disciplines Theoriques (Fle) Aux Etudiants Des Universites. *XLinguae*, 12(1), 161–169. <https://doi.org/10.18355/XL.2019.12.01XL.12>.
- Abusa'aleek, R. A., & Baniabdelrahman, A. A. (2020). The Effect of Gamification on Jordanian EFL Sixth Grade Students' Reading Comprehension. *International Journal of Education and Training (InjET)*, 6(1), 1–11. <https://www.researchgate.net/profile/Abdallah-Baniabdelrahman-2/publication/343255275>.
- Adhani, A., Basnawati, B., Zulfadli, Z., Ariesta, C. M., Siska, M., Suhaya, S., & Ramadhan, T. P. (2020). Science Literacy Capability Profile Of Prospective Biology Teachers At The University Of Borneo Tarakan. *Biopedagogia*, 2(2). <https://doi.org/10.35334/biopedagogia.v2i2.1818>.
- Batlolona, J., Baskar, C., Kurnaz, M., & Leasa, M. (2018). The Improvement of Problem-Solving Skills and Physics Concept Mastery on Temperature and Heat Topic. *Jurnal Pendidikan IPA Indonesia*, 7(3), 273–279. <https://doi.org/10.15294/jpii.v7i3.12432>.
- Bybee, R., & McCrae, B. (2011). Scientific Literacy and Student Attitudes: Perspectives From Pisa 2006 Science. *International Journal of Science Education*, 33(1). <https://doi.org/10.1080/09500693.2010.518644>.
- 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>.
- Clarke, T. (2020). Children's Wellbeing and Their Academic Achievement: The Dangerous Discourse of

- 'Trade-Offs' in Education. *Theory and Research in Education*, 18(3), 263–294. <https://doi.org/10.1177/1477878520980197>.
- Cole, J., & Feng, J. (2015). Effective strategies for improving writing skills of elementary English language learners. *Chinese American Educational Research and Development Association Annual Conference*, 1–25. <https://eric.ed.gov/?id=ED556123>.
- Elmahdi, I., Al-Hattami, A., & Fawzi, H. (2018). Using Technology for Formative Assessment to Improve Students' Learning. *Turkish Online Journal of Educational Technology-TOJET*, 17(2), 182–188. <https://eric.ed.gov/?id=EJ1176157>.
- Firmansyah, & Gradini, E. (2020). National Examination Absorption Profile For SMP/MTs Mathematics Subjects In 2019 In Central Aceh District. *Jurnal As-Salam*, 4(2). <https://doi.org/10.37249/as-salam.v4i2.184>.
- Friatma, & Anhar. (2019). Analysis of Validity, Reliability, Discrimination, Difficulty and Distraction Effectiveness In Learning Assessment. *Journal of Physics: Conference Series*, 1387(1). <https://doi.org/10.1088/1742-6596/1387/1/012063>.
- Gormally, C., Brickman, P., & Lutz, M. (2012). Developing a Test of Scientific Literacy Skills (TOSLS): Measuring Undergraduates' Evaluation of Scientific Information and Arguments. *CBE Life Sciences Education*, 11(4), 364–377. <https://doi.org/10.1187/cbe.12-03-0026>.
- Hakim, L., Lubis, P. H. M., Lefudin, L., & Sulistyowati, R. (2022). The Relationship between Learning Style toward Mastery of Concepts and Problem Solving Skill Introduction to Nuclear Physics for Pre-Service Teacher. *Jurnal Penelitian Pendidikan IPA*, 8(2), 666–673. <https://doi.org/10.29303/jppipa.v8i2.1176>.
- Holbrook, J., & Rannikmae, M. (2009). The Meaning of Scientific Management. *The Contributions of Alexander Hamilton Church to Accounting and Management*, 4(3), 275–288. <https://doi.org/10.4324/9781003056584-3>.
- Izhar, G., Wardani, K., & Nugraha, N. K. (2022). The Development Enviromental Literacy Media Learning for Elementary School Student. *Journal of Innovation in Educational and Cultural Research*, 3(3), 397–404. <https://doi.org/10.46843/jiecr.v3i3.116>.
- Jufrida, J., Basuki, F. R., Kurniawan, W., Pangestu, M. D., & Fitaloka, O. (2019). Scientific Literacy and Science Learning Achievement at Junior High School. *International Journal of Evaluation and Research in Education*, 8(4), 630–636. <https://doi.org/10.11591/ijere.v8i4.20312>.
- Kelana, J. B., Wardani, D. S., & Wulandari, M. A. (2021). Learning Methods and Critical Thinking Ability on Science Learning Outcomes. *Jurnal Ilmiah Sekolah Dasar*, 5(1), 69–76. <https://doi.org/10.23887/jisd.v5i1.29940>.
- Kibirige, I., & Teffo, W. L. (2014). Actual and Ideal Assessment Practices in South African Natural Sciences Classrooms. *International Journal of Educational Sciences*, 6(3), 509–519. <https://doi.org/10.31901/24566322.2014/06.03.1>.
- Lederman, N. G., Lederman, J. S., & Antink, A. (2013). Nature of Science and Scientific Inquiry as Contexts for the Learning of Science and Achievement of Scientific Literacy. *International Journal of Education in Mathematics Science and Technology (IJEMST) International Journal of Education in Mathematics Science and Technology International Journal of Education in Mathematics Science and Technology*, 1(3), 138–147. <https://eric.ed.gov/?id=ED543992>.
- Martinah, A. A., Mubarak, V., Miarsyah, M., & Ristanto, R. H. (2022). Development of Contextual-Based Science Literacy Test Instruments on Environmental Pollution Materials. *Bioedusiana: Jurnal Pendidikan Biologi*, 6(2), 192–218. <https://doi.org/10.37058/bioed.v6i2.3251>.
- N.M.S. Yantiningasih, I.W. Suastra, & D.B. Sanjaya. (2022). Development of Scientific Literacy Instruments and Learning Outcomes for Class VI Elementary School Students. *PENDASI Jurnal Pendidikan Dasar Indonesia*, 6(1), 13–23. [https://doi.org/10.23887/jurnal\\_pendas.v6i1.531](https://doi.org/10.23887/jurnal_pendas.v6i1.531).
- Novita, R., & Putra, M. (2016). Using Task Like PISA's Problem to Support Student's Creativity in Mathematics. *Journal on Mathematics Education*, 7(1), 31–4. <https://doi.org/10.22342/jme.7.1.2815.31-42>.
- Özer, M., & Bilgisi Öz, M. (2020). What Does PISA Tell Us About Performance of Education Systems? PISA Eğitim Sistemlerinin Performansı Hakkında Bize Ne Söylüyor? *Bartın University Journal of Faculty of Education*, 9(2), 217–228. <https://doi.org/10.14686/buefad.697153>.
- Prasetia, I., & Adlan, M. (2022). Management of the Literacy Movement Program (LMP) to Improve Reading Culture in Elementary Schools. *Journal of Innovation in Educational and Cultural Research*, 3(3). <https://doi.org/10.46843/jiecr.v3i3.117>.
- Pratiwi, S. N., Cari, C., & Aminah, N. S. (2019). 21st Century Science Learning with Students' Scientific Literacy. *Jurnal Materi Dan Pembelajaran Fisika*, 9(1), 34–42. <https://doi.org/https://doi.org/10.20961/jmpf.v9i1.31612>.

- Rahmadani, Y., Fitakurahmah, N., Funky, N., Prihatin, R., Majid, Q., & Prayitno, B. A. (2018). Profil Keterampilan Literasi Sains Siswa di Salah Satu Sekolah Swasta di Karanganyar. *Jurnal Pendidikan Biologi*, 7(3), 183. <https://doi.org/10.24114/jpb.v7i3.10123>.
- Ramadhan, W., Malahati, F., Romadhon, K., & Ramadhan, S. (2023). Analysis of Multiple Choice Questions Type Items in Elementary School Daily Assessments. *Tarbiyah Wa Ta'lim: Jurnal Penelitian Pendidikan Dan Pembelajaran*, 10(2), 93–105. <https://doi.org/10.21093/twt.v10i2.6155>.
- Richey, R. C., & Klein, J. D. (2014). Design and Development Research. In *Handbook of research on educational communications and technology* (pp. 141–150). Springer. [https://doi.org/https://doi.org/10.1007/978-1-4614-3185-5\\_12](https://doi.org/https://doi.org/10.1007/978-1-4614-3185-5_12).
- Rubini, B., Permanasari, A., & Yuningsih, W. (2018). Learning Multimedia Based on Science Literacy on the Lightning Theme. *Jurnal Penelitian Dan Pembelajaran IPA*, 4(2), 89–104. <https://doi.org/10.30870/jppi.v4i2.3926>.
- Saepuzaman, D., Istiyono, E., & Haryanto. (2022). Characteristics of Fundamental Physics Higher-order Thinking Skills Test Using Item Response Theory Analysis. *Pegem Egitim ve Ogretim Dergisi*, 12(4), 269–279. <https://doi.org/10.47750/pegegog.12.04.28>.
- Saputra, B., & Sujarwanta, A. (2021). Tranformasi Pembelajaran Berbasis Proyek Science, Technology, Engineering and Mathematics di Masa Pandemi Covid-19. *BIOLOVA*, 2(1), 1–8. <https://doi.org/10.24127/biolova.v2i1.491>.
- Sarea, M. S., & Ruslan, R. (2019). Characteristics Of Question Items: Classical Test Theory Vs Item Response Theory? *Didaktika : Jurnal Kependidikan*, 13(1). <https://doi.org/10.30863/didaktika.v13i1.296>.
- Setyawarno, D. (2017). *Effort To Improve The Quality Of Qualifications With Quest Application Analysis*.
- Siahaan, S. R., Situngkir, D., Nitami, M., & Wahidin, M. (2022). The Relationship Between Job Insecurity and Job Stress and Employee Turnover Intention. *Lontara Journal of Health Science and Technology*, 3(2). <https://doi.org/10.53861/lontarariset.v3i2.304>.
- Singh, B. (2019). Character Education in the 21st Century. *Journal of Social Studies (JSS)*, 15(1), 1–12. <https://doi.org/10.21831/jss.v15i1.25226>.
- Suharsimi Arikunto. (2021). *Dasar-Dasar Evaluasi Pendidikan Edisi 3* (Restu Damayanti (ed.); Ketiga). PT Bumi Aksara.
- Sulaiman, S., Radzi, M. Q. A.-N. A., Kamaruddin, N., & Ali, W. A. F. W. (2020). *Typography Usage in Multimedia Teaching Aid Interface Design*. 502(Imdes), 157–162. <https://doi.org/10.2991/assehr.k.201202.070>.