Birlesik Dünya Arastırma BD-CENTER Innovasyon ve Yayıncılık Merkezi

Cypriot Journal of Educational Sciences

Volume 17Issue 12 (2022) 4361–4378

CJES Cypriot Journal of Educational Sciences

www.cjes.eu

Developing testing instruments to measure science literacy of elementary school students

Ketut Susiani [1]**,** Universitas Pendidikan Ganesha, Singaraja 81116, Indonesia https://orcid.org/0000-0002-2530-4946

Nyoman Dantes, Universitas Pendidikan Ganesha, Singaraja 81116, Indonesia https://orcid.org/0000-0001-8025-4243

Ida Bagus Putu Arnyana, Universitas Pendidikan Ganesha, Singaraja 81116, Indonesia https://orcid.org/0000-0002-0385-0065

Ni Ketut Suarni, Universitas Pendidikan Ganesha, Singaraja 81116, Indonesia https://orcid.org/0000-0003-0112-5123

I Kadek Suartama, Universitas Pendidikan Ganesha, Singaraja 81116, Indonesia https://orcid.org/0000-0002-4881-3233

Kadek Suranata, Universitas Pendidikan Ganesha, Singaraja 81116, Indonesia https://orcid.org/0000-0002-1389-5363

Suggested Citation:

Susiani, K., Dantes, N., Arnyana, I. B. P., & Suarni, N. K., Suartama, I. K., & Suranata, K. (2022). Developing testing instruments to measure science literacy of elementary school students. *Cypriot Journal of Educational Science*. *17*(12), 4361–4378. https://doi.org/10.18844/cjes.v17i12.7730

Received from August 15, 2022; revised from October 11, 2022; accepted from December 20, 2022.

Abstract

Literacy is one aspect of knowledge that has become a benchmark for the success of education. The measurement of the level of quality of education carried out by UNESCO is recorded at the lowest rank. The purpose of this study was to develop test instruments to measure the scientific literacy abilities of elementary school students. The number of respondents in this study was 20 students and 12 items. This research method used the design of (1) a preliminary study stage, (2) a development study stage and (3) an evaluation stage. In the first stage, the preliminary study was reviewed through literature studies, field studies and identification according to the needs of the instrument. Furthermore, the second stage of the development study was carried out by designing the initial product, validation tests by experts, analysis and revisions, and then carrying out initial trials, analysis and refinement. Data analysis was carried out using the Rasch model approach. The results of the analysis showed the reliability index of the instrument (=0.97) and the reliability of the person (α = 0.81). In general, this instrument can explain 72.7% of the variance in respondents, so this test instrument can be used to measure the scientific literacy of elementary school students.

Keywords: Elementary school, instrument test, Rasch model, science literacy, student

[1] *ADDRESS FOR CORRESPONDENCE: Ketut Susiani, Universitas Pendidikan Ganesha, Singaraja 81116, Indonesia.
E-mail address: ketut.susiani@undiksha.ac.id, susiani@undiksha.ac.id / Tel.: +62 819-3657-0559

1. Introduction

At this time, and especially in the future, the existence, position and role of knowledge have become strategic and main things (Ahmad & Karim, 2019). The future is determined by knowledge, so that the world joins and rests on knowledge (Bolisani & Bratianu, 2017). Knowledge is the most valuable and most needed aspect. Without knowledge, people (even nations and countries) will be marginalised and left behind. On the contrary, with good knowledge, people, nations and countries can become winners in various life activities. And the knowledge needed and suitable for the future can be known by looking at the trends in knowledge change that lead to the future. While in the aspect of students, many changes have happened to them because of technological changes that are always presented to them every day and even every time. These changes, according to Curley and Salmelin (2018), are like control. They are people who like group ties and social ties; it is just that they build groups through their social media, and therefore, their groups cross nations, countries, cultures and even religion; they also like having a lot of choices and are open-minded people. Seeing the changes that have occurred, it is necessary to have a competency that must be developed in the future of elementary school students, namely scientific literacy competence.

1.1. Theoretical and conceptual framework

In the context of Programme for International Student Assessment (PISA), scientific literacy is defined as the ability to use scientific knowledge, identify questions and draw conclusions based on evidence in order to understand and make decisions regarding nature and changes made to nature through human activities (Pahrudin, 2018, 2019; Rohmawati et al., 2018; Sharon & Baram-Tsabari, 2020). This definition of scientific literacy views scientific literacy as multidimensional, which is not just an understanding of scientific knowledge but broader than that. PISA 2000 and 2003 define three major dimensions of scientific literacy in their measurement, namely competence/science process, content/knowledge of science and context of science application (Nuril, 2020). In PISA 2006, the dimensions of scientific literacy were developed into four dimensions, the addition of which is the aspect of students' attitudes towards science (She et al., 2018).

The context aspect of PISA assesses scientific knowledge as relevant to the science education curriculum in the participating countries without limiting itself to general aspects of the national curriculum of each country. The PISA assessment is framed within a broader general life situation and is not limited to life at school. The items in the PISA assessment focus on situations related to individuals, families and individual groups (personal), community (social) and cross-country (global) life (Salas-Velasco et al., 2021). The context of PISA includes areas of application of science in personal, social and global settings, namely (1) health; (2) natural resources; (3) environmental quality; (4) danger; and (5) the latest developments in science and technology.

The results of the 2018 PISA survey, published in March 2019, photographed some of Indonesia's education problems. In the category of reading, science and mathematics, Indonesia's score is low because it is ranked 74th out of 79 countries. This can be seen in Figure 1.

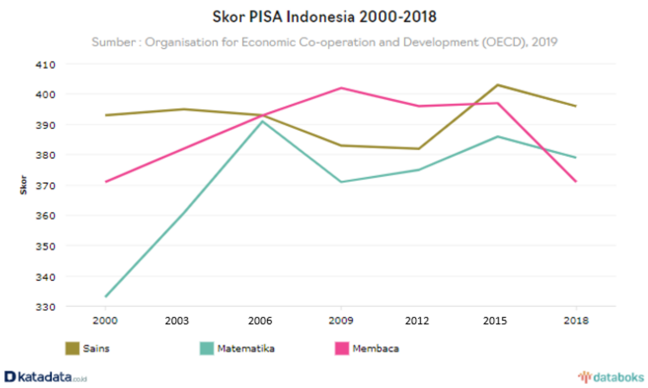

Figure 1: *Indonesia's PISA Scores From 2000 to 2018*

The PISA scores in Figure 1 indicate that Indonesia is still weak in these three aspects. If we look at the graph, science from 2000 had increased to 405 in 2015 and decreased again in 2018 (Nugrahanto & Zuchdi, 2019). PISA is an evaluation survey of the world's education system that measures the performance of middle school students (Mahmut, 2020). This assessment is conducted every 3 years and is divided into three main points, namely literacy, mathematics and science. The results in 2018 measured the abilities of 600,000 15-year-olds from 79 countries (Chang & Bangsri, 2020). Previously, the Organisation for Economic Cooperation and Development (OECD) announced the results of the 2018 PISA. As in previous years, Indonesia's ranking was not satisfactory. According to data published by the OECD (2019), from the 2009 to 2015 survey period, Indonesia consistently ranks in the bottom 10 (Henukh et al., 2021).

1.2. Related research

Based on a report by the *Association for Evaluation of Educational Achievement International*, it is reported that the quality of Indonesian students is in 36 out of 50 countries. In the fields of mathematics and reading skills, it is ranked 39th out of 41 countries surveyed (Han et al., 2016;). More specifically, student achievement in science, a test conducted by *Trends International Mathematics and Sciences Study* (TIMSS), an institution that measures educational outcomes in the world, reports that the science ability of Indonesian elementary school students is ranked 32nd out of 38 countries (Aviyanti, 2020; Kartianom & Retnawati, 2018; Tze & Li, 2021). The *PISA* report showed that of the 41 countries surveyed, for the field of science, Indonesia, was ranked 38th (Kembara et al., 2020). These facts show that the quality of science learning in Indonesia, including Bali Province, needs to be improved because science learning plays an important role in improving the quality of human resources (Sari et al., 2021).

The results of Warpala's (2006) research revealed that the science learning achievement of elementary school students in Singaraja ranked sixth out of seven subjects (Pancasila, religion, Indonesian, social studies, English, science and mathematics). In relation to learning science, it is important to instil aspects of understanding in students. This is in line with the constructivist view, which states that understanding is the most fundamental element in learning (ELIBRARY, n.d.; Fauth et al., 2019; Supena et al., 2021). Gardner (1999) defines understanding as a mental process of adaptation and transformation of knowledge. Furthermore, it is said that someone is said to understand something if he is able to show that understanding in the context of the same or different understanding. This can be seen based on their ability to communicate the ideas they have and be able to solve the problems they face. Thus, problem-solving skills require an understanding of one's self to the problems at hand.

Farina (2019) reveals that there are at least three factors that are the main barriers for students to achieving understanding. These three factors are (1) the learning methods applied tend to tolerate *unitary ways of knowing*; (2) the curriculum substance tends to be decontextual; and (3) the formulation of goals is rarely oriented towards achieving deep understanding. Besides being able to hinder understanding, these technical and substantial errors can cause misunderstandings in students, which are known as misconceptions.

The problem of misconceptions can happen to anyone, students, teachers, textbooks or learning methods (Moodley & Gaigher, 2019; Mufit et al., 2020). Misconceptions that occur in students continuously can interfere with the formation of scientific conceptions (concept understanding) (Widiyatmoko & Shimizu, 2018a, 2018b). The next problem faced by the teacher is how to change the knowledge of students who have still labelled misconceptions as scientific conceptions (Nadelson et al., 2018). Besides misconceptions being a problem in the learning process, there are other things that happen at school. Judging from the real conditions in schools, phenomena that we can encounter in schools such as the low motivation of students to complete their assignments on time, like spending more time, noting want to finish work till completed, being less focused and tending to feel lazy.

Other symptoms that appear are the lack of students' scientific literacy in solving problems that arise during the learning process activities and the weak enthusiasm of students to be active in class (Johnson, 2021). The lack of scientific literacy and the weak enthusiasm of students to be active in class will be seen from the attitude of students who tend to be passive, only listening without asking if something is not understood and giving up easily when unable to answer questions or assignments given by the teacher (Santyasa et al., 2020).

Lack of Scientific literacy will be indicated by their inability to complete the work and tasks given properly to completion, lack of self-effort, work carelessly and easily giving up when doing tasks that are considered quite hard (Indrawati & Wardono, 2019; Wijayani, 2020). Meanwhile, students with high scientific literacy will show the opposite, such as being able to complete the work and assignments given properly to the end, their own enthusiasm and hard work, working with full responsibility and always trying to find a way out of any problems found when doing the tasks and work given (Heksa, 2021; Wen et al., 2020).

The scope of science literacy for elementary school age children is to (1) stimulate students to be ready to learn; (2) involve students in learning; and (3) create a fun learning atmosphere. The above is one way to create a pleasant learning atmosphere. Scientific literacy makes it easier for students to adapt to the progress of science and technology that continues to develop and can stimulate students' imagination and creativity (Efendi, 2021).

A new view of scientific literacy (Yuliati, 2017) explains that scientific literacy is the ability to use scientific knowledge to identify problems and how to draw conclusions from evidence in order to understand and make decisions about nature and the changes that occur. Nature's activities are carried out by humans through scientific literacy, which is important in today's 21st century. On the other hand, Effectiveness et al. (2020) explain that scientific literacy increases with the help of gadget-based interactive multimedia.

On the other hand, science subject teachers also have difficulty figuring out how to make students have the motivation and responsibility to do the tasks given. This is one of the foundations for developing a test instrument to measure students' scientific literacy in science subjects.

1.3. Purpose of the study

The purpose of this research is to develop a scientific literacy test instrument for elementary schools.

2. Method and materials

This section will explain the research methods used, including the research sample or participants, the data analysis used and the data collection process.

2.1. Research model

Research on the development of a test instrument to measure scientific literacy used a research design (Firdaus & Purnama, 2018), where the design has been simplified into three main stages; (1) preliminary study stage; (2) development study stage; and (3) evaluation stage.

2.2. Participants

This research is a limited test study using 20 students from different schools in Denpasar City. All the students come from a high class, namely the fifth grade of elementary school.

2.3. Data collection tools

The data collection tool uses a questionnaire distributed via Google Form, considering the pandemic situation, when the research was conducted. The questionnaire used is in Indonesian.

2.4. Data collection process

Data collection is carried out in accordance with the stages of the research design. In the first stage, the preliminary study was reviewed through literature studies, field studies and identification according to the needs of the instrument. Furthermore, in the second stage, the development study was carried out by designing the initial product, validation tests by experts, analysis and revisions; then, preliminary trials, analysis and refinements were carried out to produce a product in the form of a hypothetical instrument. In addition to the three stages, a characteristic test was also carried out with the aim of determining the level of difficulty, distinguishing power and the proportion or comparison of scientific literacy categories in the instrument.

2.5. Data analysis

The instrument developed was used to measure students' scientific literacy skills. Data analysis was carried out using the Rasch model and assisted by the developed Winstep software (Linacre, 2006, 2008). The analysis was carried out with data sourced from 20 high-class student respondents and an instrument consisting of a total of 12 items. Data tabulated in MS Excel was then converted and analysed with Winstep 3.73 software in the Windows 7 operating system.

3. Results

3.1. Instrument reliability

The instrument reliability analysis was carried out using the chi-square test, with a score of 408.85 and a degree of freedom (df) of 206 (p = 0.00 and p < 0.01). This shows that the overall measurements made are very good and the results are significant. The results of this analysis contain two outputs, namely the output for the respondent (person) and the output for the item. The respondent table describes the general fit or not of the respondents used. Likewise, the item table explains whether in general the items used in the instrument can be said to be fit or not; refers to Table 1. This shows that, in general, respondents have sufficient non-cognitive scores and instrument items are said to be good.

Referring to Table 2, the mean value of the measurements obtained in the person test is 1.24 (μ > 0.00). This shows that, in general, respondents have high scientific literacy scores, in the sense that respondents have a tendency to agree on items that measure scientific literacy indicators. The logit value of 1.24 also indicates that the respondents do not have too large a diversity in the measured constructs. This happens because the respondents come from the same demographic settings. The SEPARATION index in the respondent's table shows a value of 2.09. With the SEPARATION index = 2.09, the strata of the respondents in this study can be seen using the person strata formula (Misbach & Sumintono, 2014) as follows:

$$H = \frac{[(4 \: x \: SEPARATION) + 1]}{3}$$

where,

H: Person strata

SEPARATION: Value for respondents generated

Based on the formula, the *H* value obtained is 3, 12. This shows that the respondents can be divided into three major groups: high, medium and low. Based on the SEPARATION index in the item table, the value of the strata item is obtained based on the same formula as the person strata, which is = 7.72. This indicates that the items in this test are able to classify students' abilities in scientific literacy up to eight criteria. This shows that these test items are able to accurately assess respondents' answers related to the construct of scientific literacy. Table 1 shows a summary of 20 measured person.

Table 1

Summary of Instrument Statistics: Reliability of Respondents and Items Summary of 113 Measured Persons

	Total score	Count	Measure	Model error	Infit MNSQ	ZSTD	Outfit MNSQ	ZSTD
Mean	41.3	12.0	1.24	0.50	0.94	−0.1	1.00	0.0
SD	4.8	0.0	1.22	0.04	0.32	0.8	0.36	0.8
Max.	53.0	12.0	4.57	0.65	1.74	1.7	1.78	1.8
Min.	34.0	12.0	−0.52	0.48	0.56	−1.2	0.59	−1.1
Real RMSE 0.53 True SD 1.10 SEPARATION 2.09 Person reliability 0.81								
Model RMSE 0.50 True SD 1.12 SEPARATION 2.23 Person reliability 0.83								
SE of person mean = 0.28								

Person raw score-to-measure correlation = 1.00.

Cronbach's alpha (KR-20) person raw score 'test' reliability = 0.81.

Table 2

Summary of 12 Measured Items

	Total score	Count	Measure	Model error	Infit MNSQ	ZSTD	Outfit MNSQ	ZSTD
Mean	68.9	20.0	0.00	0.39	0.97	−0.1	1.00	−0.1
SD	16.3	0.0	2.34	0.03	0.38	1.1	0.45	1.3
Max.	89.0	20.0	3.50	0.44	1.99	2.7	2.16	2.9
Min.	44.0	20.0	−3.11	0.35	0.65	−1.2	0.62	−1.3
Real RMSE 0.42 True SD 2.30 SEPARATION 5.54 Item reliability 0.97								
Model RMSE 0.39 True SD 2.31 SEPARATION 5.95 Item reliability 0.97								

U Mean= 0.0000; *U* Scale= 1.0000.

Item raw score-to-measure correlation= −1.00.

240 data points. Log-likelihood chi-square: 408.85 with d.f. and p = 0.0000;

Global root-mean-square residual (excluding extreme score) = 0.5640.

Cronbach's alpha (KR-20), which measures the interaction between respondents and items, shows good results, with alpha = 0.81. The reliability value for the respondents obtained based on Table 1 is 0.81. This shows that there is a match between the respondents and the instruments used. In addition, the reliability value for the item is 0.97, which indicates that the instrument has very good reliability (Sumintono & Widhiarso, 2014). Based on the evaluation of the psychometric properties, it can be said that overall the actual data obtained are in accordance with the requirements of the Rasch model so that further analysis can be applied.

3.2. Validity

In the analysis using the Rasch model, the interpretation of measurements, especially content validity and construct validity, can be evaluated more precisely. In addition, researchers can also estimate the validity of the respondents by looking at the respondents who have the most inconsistent answers. Figures 1 and 2 show the interaction between respondents and items based on the variables of gender and perceptions of the test instrument. In Figure 1, l is male and p is female. In the box in Figure 2, X is the respondent who perceives that students are able to answer the test well. Based on the two figures, information can be obtained that most of the respondents have a high scientific literacy level. Referring to the perception variable on the application of test instrument values in gender, there is no significant difference at the level of scientific literacy.

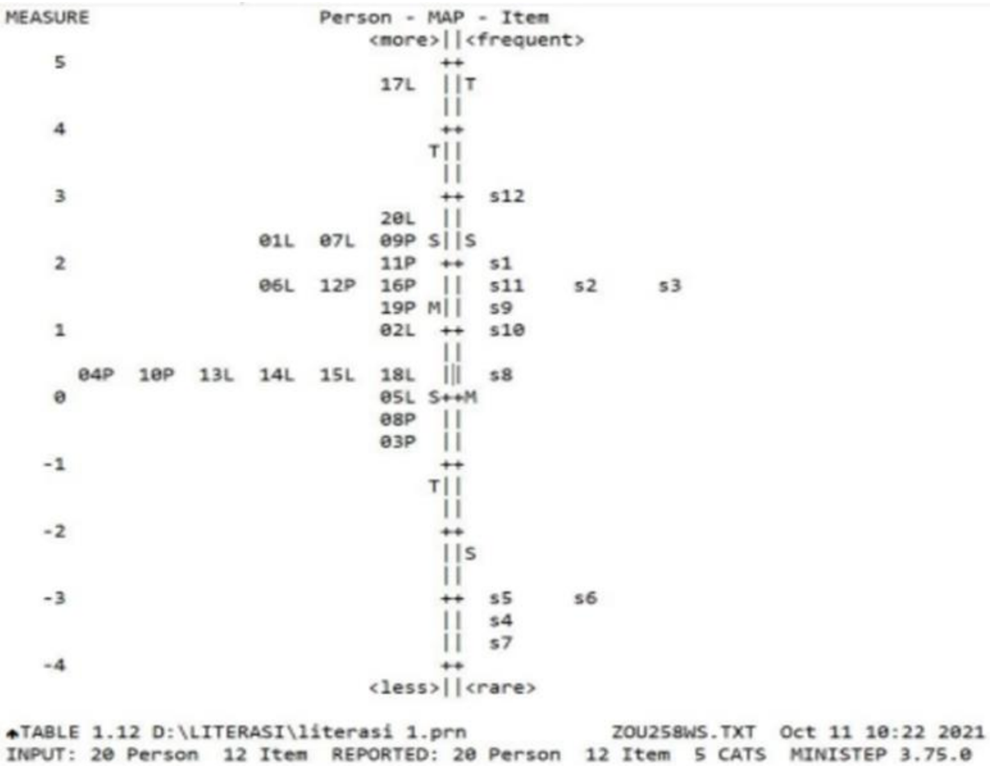

Figure 2

Distribution Map of Respondents and Items in the Logit Bar Based on Gender

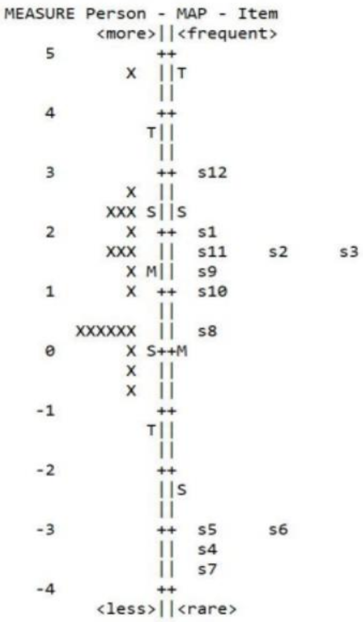

Figure 3

Respondent's Distribution Map and Logit Cross Item Based on the Perception of Enthusiasm in Answering the Questions

Based on Figure 3, it can be found that the item that is difficult to be answered by the respondents is item S12, which states 'why is sitting in the train said to be moving towards the station but is said to be motionless towards the friend beside you?' In addition, an item that was too easy for the respondents was also found. This item is S8. It is possible that this item contains a bias in everyday life so that respondents tend to answer correctly and scientifically.

In addition, information was also obtained that 4 items were below logit 0. Based on the logit rule, it was also obtained that most of the respondents were at the intermediate level. At the sub-dimensional level, it can be said that the sub-dimension that is perceived to be the easiest to answer or the best answer to the respondent is the sub-dimension of knowledge that the nature of light propagates in a straight line. Referring to the distribution map of the respondents, there was no difference in level, either based on the gender variable or perception of the science test. This shows that the symptoms of the scientific literacy test instrument can develop in both male and female groups. In addition, scientific literacy test instruments can develop on both tests that are categorised as easy and tests with difficult categories.

3.3. Validity of respondents and items

As an attempt to examine respondents and items that do not fit (outliers or misfits), Sumintono and Widhiarso (2013) suggest the following three criteria:

1. Outfit mean square score (MNSQ) used is 0.5 < MNSQ < 1.5;
2. Outfit Z-Standard score (ZSTD) used is −2.0 < ZSTD < +2.0;
3. Point measure correlation score (Pt Mean Corr) used is 0.4 < Pt Measure Corr < 0.85.

Table 3

Result of Fit/Misfit Respondent Test

Entry number	Total score	Total count	Measure	Model SE	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD	PT-measure CORR	PT-measure EXP	Exact OBS %	Match EXP%	Person
9	46	12	2.33	0.51	1.74	1.7	1.78	1.8	A 0.40	0.81	58.3	61.1	09P
17	53	12	4.57	0.65	1.23	0.6	1.77	1.1	B 0.64	0.80	66.7	74.3	17L
18	37	12	0.18	0.48	1.59	1.4	1.61	1.5	C 0.89	0.82	41.7	59.1	18L
14	38	12	0.41	0.48	1.46	1.2	1.53	1.3	D 0.90	0.82	41.7	60.3	14L
15	38	12	0.41	0.48	1.05	0.3	1.10	0.4	E 0.88	0.82	58.3	60.3	15L
4	38	12	0.41	0.48	1.00	0.1	1.04	0.2	F 0.89	0.82	58.3	60.3	04P
19	42	12	1.34	0.49	0.96	0.0	0.98	0.1	G 0.73	0.81	41.7	59.3	19P
5	36	12	−0.05	0.48	0.94	0.0	0.96	0.0	H 0.84	0.82	66.7	57.5	05L
3	34	12	−0.52	0.49	0.89	−0.2	0.89	−0.2	I 0.80	0.83	58.3	54.6	03P
10	37	12	0.18	0.48	0.83	−0.3	0.88	−0.2	J 0.87	0.82	58.3	59.1	10P

6	43	12	1.58	0.49	0.87	−0.2	0.79	−0.5	J 0.89	0.81	66.7	60.1	06L
9	35	12	−0.28	0.48	0.81	−0.4	0.84	−0.3	i 0.86	0.83	50.0	55.5	08P
12	43	12	1.58	0.49	0.79	−0.4	0.82	−0.3	h 0.86	0.81	66.7	60.1	12P
11	45	12	2.07	0.51	0.78	−0.5	0.80	−0.4	g 0.78	0.81	83.3	61.5	11P
13	38	12	0.41	0.48	0.72	−0.7	0.80	−0.4	f 0.85	0.82	75.0	60.3	13L
2	41	12	1.10	0.48	0.75	−0.6	0.76	−0.5	e 0.82	0.81	58.3	59.9	02L
7	46	12	2.33	0.51	0.73	−0.6	0.71	−0.7	d 0.91	0.81	58.3	61.1	07L
20	47	12	2.60	0.52	0.56	−1.2	0.73	−0.6	c 0.84	0.81	58.3	60.7	20L
1	46	12	2.33	0.51	0.60	−1.0	0.60	−1.1	b 0.87	0.81	75.0	61.1	01L
16	44	12	1.82	0.50	0.57	−1.1	0.59	−1.1	a 0.83	0.81	66.7	60.8	16P

Referring to Table 3, of the 20 research respondents, all answered well. So, in the context of analysis with inferential statistics, it is recommended that respondents need not be eliminated. For item fit/misfit analysis, three criteria are still used as previously stated. However, the criteria for the elimination of items are based on the results of the analysis that are absolutely convincing that the items are very consistent. Based on Table 3, the average logit value of the items is 0.0. This shows that, as a whole, the instrument is able to measure what it is meant to measure. The average value of the item 0.0 logit is a random value that is set to express the probability of 50:50 as an equivalent measure between the respondent's level of reliability and the difficulty of the item (Heene, 2020).

Table 4

Item Fit/Misfit Test Results

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
11	81	20	-1.71	.40	1.99	2.7	2.16	2.9	A .10	.54	45.0	61.3	s11
1	82	20	-1.88	.40	1.38	1.2	1.63	1.8	B .18	.53	50.0	61.8	s1
3	81	20	-1.71	.40	1.14	.5	1.10	.4	C .71	.54	45.0	61.3	s3
12	89	20	-3.11	.44	1.10	.4	1.13	.4	D .22	.48	55.0	68.0	s12
10	76	20	-.93	.39	.98	.0	.98	.0	E .56	.56	60.0	61.3	s10
2	81	20	-1.71	.40	.88	-.3	.84	-.4	F .63	.54	65.0	61.3	s2
7	44	20	3.50	.35	.82	-.6	.81	-.6	f .72	.61	55.0	54.1	s7
8	72	20	-.31	.39	.68	-1.0	.69	-1.0	e .58	.57	70.0	63.3	s8
5	48	20	3.00	.35	.68	-1.1	.68	-1.1	d .84	.61	70.0	57.5	s5
6	48	20	3.00	.35	.65	-1.2	.68	-1.1	c .79	.61	60.0	57.5	s6
4	46	20	3.25	.35	.67	-1.2	.66	-1.2	b .76	.61	70.0	55.2	s4
9	79	20	-1.39	.40	.65	-1.2	.62	-1.3	a .51	.55	80.0	61.5	s9
MEAN	68.9	20.0	.00	.39	.97	-.1	1.00	-.1			60.4	60.3	
S.D.	16.3	.0	2.34	.03	.38	1.1	.45	1.3			10.5	3.6	

TABLE 10.3 D:\LITERASI\literasi 1.prn ZOU258WS.TXT Oct 11 10:22 2021
INPUT: 20 Person 12 Item REPORTED: 20 Person 12 Item 5 CATS MINISTEP 3.75.0

Based on Table 4, it is obtained that item C11 has an outfit mean square (MNSQ) value of 2.16 > 1.5 and a point measure correlation value of 0.10. This indicates that the item is a misfit and so it is recommended to be eliminated. However, based on the bubble chart item, C11 can still be used because it does not come out of the mean square outfit. Then, judging from the quality of the test instrument item C12, which is classified as the most difficult to analyse, it has an outfit mean square (MNSQ) value of 1.13 < 1.5 and a point mean correlation value of 0.22. This identifies that item C12 is very fit. Figure 4 shows the results of the bubble chat.

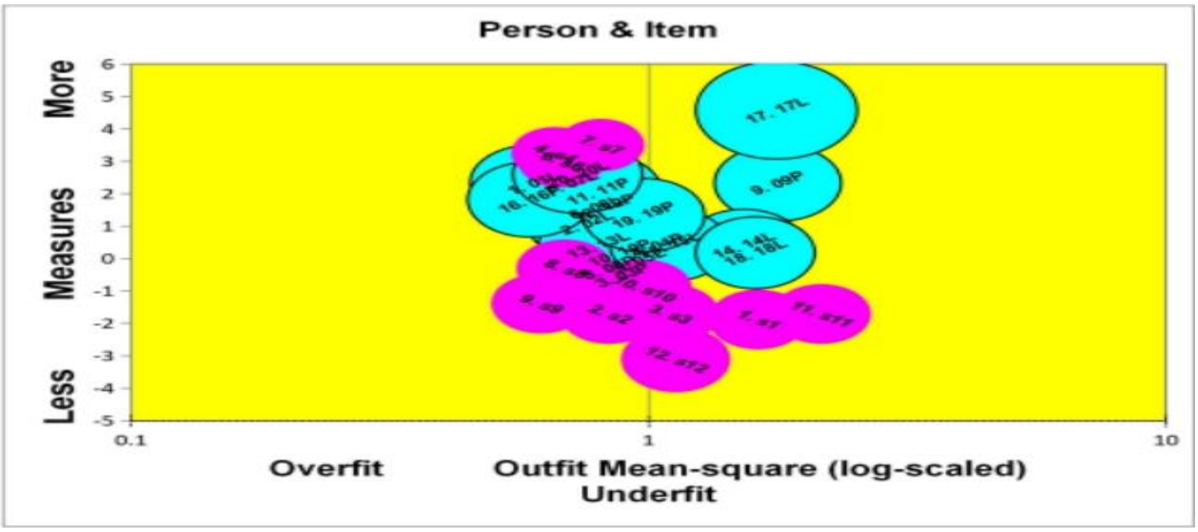

Figure 4

The Results of the Bubble Chat

3.4. Unidimensionality of the instrument

In this case, the student's scientific literacy construct unidimensionality is an important measure to evaluate whether the developed instrument is able to measure what it is supposed to measure. The analysis of the Rasch model uses principal component analysis of the residuals, which is to measure the extent to which the diversity of the instruments measures what should be measured (Sumintono & Widhiarso, 2014).

Table 5

Unidimensionality Instrument Test Results

		Empirical	Modelled	
Total raw variance in observations	43.9	100.0%	100.00%	
Raw variance explained by measures	31.9	72.7%	71.1%	
Raw variance explained by persons	6.0	13.7%	13.4%	
Raw Variance explained by items	25.9	58.9%	57.6%	
Raw unexplained variance (total)	12.0	27.3%	100.0%	28.9%
Unexplained variance in first contrast	2.5	5.6%	20.5%	
Unexplained variance in second contrast	1.9	4.3%	15.8%	
Unexplained variance in third contrast	1.5	3.4%	12.4%	
Unexplained variance in fourth contrast	1.3	3.1%	11.2%	
Unexplained variance in fifth contrast	1.1	2.6%	9.5%	

Based on Table 5, it can be seen that the raw data variance measurement results are 72.7%. The value is not much different from the expected value, which is 71.1%. This shows that the unidimensionality requirement of 20% can be met. In addition, the unidimensional limit in the Rasch model (Linacre as cited in Misbach & Sumintono, 2014) of 40% is also fulfilled. Another thing that is also supported is that the variances that cannot be explained by the instrument are all below 10%. This shows that the level of independence of the items in the instrument is in the good category.

3.5. Bias detection of the items

Item bias in this measurement is seen based on two variables: gender and perception of the test instrument. The analysis of the Rasch model shows the detection of item bias in differential item functioning (DIF). Bias can be identified based on the probability value of items that are below 5% (Sumintono & Widhiarso, 2009, 2014).

Table 6

Results of Bias Detection Analysis Based on Gender

Person CLASSES	SUMMARY DIF CHI-SQUARE	D.F.	PROB.	BETWEEN-CLASS MEAN-SQUARE	t=ZSTD	Item Number	Name
2	2.9972	1	.0834	1.7451	.9041	1	s1
2	.4235	1	.5152	.2270	-.3558	2	s2
2	.4235	1	.5152	.2270	-.3558	3	s3
2	.0069	1	.9338	.0037	-1.3222	4	s4
2	.6983	1	.4034	.3753	-.1197	5	s5
2	.6983	1	.4034	.3753	-.1197	6	s6
2	1.8332	1	.1758	1.0243	.4885	7	s7
2	.9528	1	.3290	.5147	.0501	8	s8
2	.0056	1	.9402	.0029	-1.3470	9	s9
2	.1340	1	.7143	.0705	-.7737	10	s10
2	1.9410	1	.1636	1.0913	.5341	11	s11
2	.3021	1	.5826	.1617	-.4943	12	s12

Table 7

Results of Bias Detection Analysis Based on the Test Instrument

Person CLASSES	SUMMARY DIF CHI-SQUARE	D.F.	PROB.	BETWEEN-CLASS MEAN-SQUARE	t=ZSTD	Item Number	Name
1	.0000	0	1.0000	.0000	.0000	1	s1
1	.0000	0	1.0000	.0000	.0000	2	s2
1	.0000	0	1.0000	.0000	.0000	3	s3
1	.0000	0	1.0000	.0000	.0000	4	s4
1	.0000	0	1.0000	.0000	.0000	5	s5
1	.0000	0	1.0000	.0000	.0000	6	s6
1	.0000	0	1.0000	.0000	.0000	7	s7
1	.0000	0	1.0000	.0000	.0000	8	s8
1	.0000	0	1.0000	.0000	.0000	9	s9
1	.0000	0	1.0000	.0000	.0000	10	s10
1	.0000	0	1.0000	.0000	.0000	11	s11
1	.0000	0	1.0000	.0000	.0000	12	s12

Referring to the results of the DIF analysis, there were no items that contained bias. It is identified based on the probability value that moves between 0.0834 and 0.338 ($p > 0.05$) for the detection of bias by gender and 1.00 ($p > 0.05$) for the detection bias based on the perception of the test instrument (Table 7). Based on the results of this analysis, it can be assumed that the items are perceived the same by respondents of different sexes and by different respondents based on the perceptions of the test instrument.

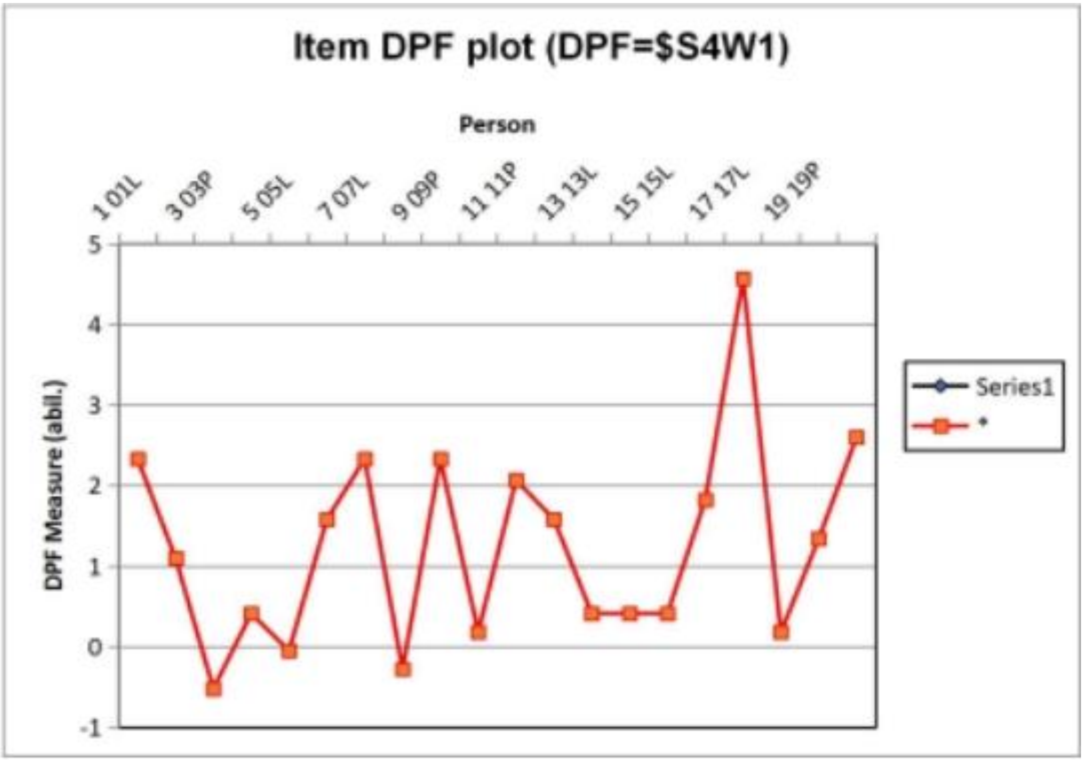

Figure 5

The Results of the DIF Analysis

3.6. Validity of the rating scale

The validity of the rating scale is a test carried out to verify whether the rating options used are confusing for respondents or not. The analysis of the Rasch model provides a verification process for the rating assumptions given in the instrument. In this instrument, five answer choices were given in the form of a Likert scale for each item. Respondents provided answers to each item given. Respondents' answers are seen based on the tendency of whether the answers move to the leftmost column (STS) or the rightmost column (SS). This choice is contradictory.

Table 8

Results of Rating Scale Validity

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	ANDRICH THRESHOLD	CATEGORY MEASURE	
1	1	12	5	-2.98	-3.00	.95	.92	NONE	(-4.94)	1
2	2	36	15	-2.59	-2.21	.49	.47	-3.75	-2.75	2
3	3	63	26	.39	.27	.71	.72	-1.71	-.23	3
4	4	91	38	2.86	2.60	1.06	1.29	1.23	2.75	4
5	5	38	16	3.73	4.21	1.43	1.40	4.24	(5.38)	5

In Table 8, it can be seen that the average observation starts from logit −2.98 for choice 1 (STS) and increases to logit 3.73 for choice 5 (SS). The increase in the logit value shows consistent results. This shows that a rating scale of 1–5 can be said to be non-confusing for respondents and is an appropriate scaling range in this instrument. Another recommended measure is the Andrich threshold to test whether the polytomy value used is correct or not. The Andrich threshold value, which moves from NONE, is then negative and leads to positive, respectively, indicating that the five options given are valid for the respondents.

3.7. Guttman's scalogram of responses

Person |Item

|1 1 1

|212319085647

|------------

17 +545545544444 17L

20 +445554443333 20L

1 +544544543332 01L

7 +555544443223 07L

9 +444444354433 09P

11 +545434443333 11P

16 +444544443233 16P

6 +554444443231 06L

12 +444454542322 12P

19 +444534433323 19P

+444454332332 02L

4 +543445332212 04P

13 +453344432222 13L

14 +534554332211

+

1444322

2

1543523443 05L

8 +534334331222 08P

3 +433333442122 03P

|------------

|1 1

|212319085647

These results indicate that there is no response that cheats or answers inconsequentially; this is due to the characteristics of students and classes who are active and conducive.

4. Discussion

Scientific literacy is one aspect of knowledge that is a benchmark for the success of an education (Aqil, 2018; Handayani, 2020; Meuthia & Ahmad, 2021). In measuring the level of quality of education carried out by UNESCO through PISA, scientific literacy is one aspect that is measured. Regarding the results of these measurements, children in Indonesia are listed in the lowest rank of student achievement in the field of science. A test that has been carried out by the TIMSS, an institution that measures educational outcomes in the world, reports that the science abilities of elementary school students Indonesia is ranked 32nd out of 38 countries (Fenanlampir et al., 2019; Kembara et al., 2020; Kristyaningrum & Winarto, 2020; Susiani et al., 2022). The 2003 PISA report showed that of the 41 countries surveyed, for the science field, Indonesia ranked 38th. These facts show that the quality of science learning in Indonesia, including Bali Province, needs to be improved, because science learning plays an important role in improving the quality of human resources (Rusyadi, 2021).

The scope of science literacy for elementary school age children is to (1) stimulate students to be ready to learn; (2) involve students in learning; and (3) create a fun learning atmosphere. The above is one way to create a pleasant learning atmosphere. Scientific literacy makes it easier for students to adapt to the progress of science and technology that continues to develop and can stimulate students' imagination and creativity (Efendi et al., 2021)

Several opinions regarding scientific literacy (e.g., Snow, 2013) state that scientific literacy means appreciation of science by increasing the learning component in oneself, so that it can contribute to the social environment. Based on the statement above, scientific literacy has a broad meaning; everyone can contribute in interpreting scientific literacy. The development of scientific literacy instruments will be able to assist teachers in developing students' potential in facing the global era.

5. Conclusion

Overall, based on the results of data analysis using the Rasch model, it can be concluded that the students' scientific literacy test developed by statistical and psychometric rules can be used as an instrument in research or assessment related to elementary school students' scientific literacy. In addition, the results of data analysis showing the reliability value of Cronbach's alpha (KR-20) of 0.81 and item reliability of 0.97 provided empirical support for the quality of measuring scientific literacy of elementary school students with this instrument. In general, the respondents have a high level of scientific literacy. Scientific literacy is indicated not to be related to gender, so it can be interpreted that the phenomenon of scientific literacy is able to develop in both male and female groups. In addition, referring to the unidimensionality of the instrument, the results of the analysis showed that the measurement is able to explain up to 72.7% of the variance that arises in the respondent group

6. Recommendation

This study only developed scientific literacy test instruments for elementary school children in Denpasar City on a limited basis. It is recommended that future research be developed with more participants and by using different analyses and items that further increase science literacy.

References

Ahmad, F., & Karim, M. (2019). Impacts of knowledge sharing: A review and directions for future research. *Journal of Workplace Learning*. https://www.emerald.com/insight/content/doi/10.1108/JWL-07-2018-0096/full/html

Aqil, D. I. (2018). Science literacy as a concept of learning biology textbooks in schools. *Didactic Discourse, 5*(02), 160–171.

Aviyanti, L. (2020). *An investigation into Indonesian pre-service physics teachers' scientific thinking and conceptual understanding of physics*. https://flex.flinders.edu.au/file/a44a1398-06d4-451f-808c-6e1a702e060b/1/AviyantiThesis2020_LibraryCopy.pdf

Bolisani, E., & Bratianu, C. (2017). Knowledge strategy planning: An integrated approach to manage uncertainty, turbulence, and dynamics. *Journal of Knowledge Management*. https://www.emerald.com/insight/content/doi/10.1108/JKM-02-2016-0071/full/html

Chang, Y. C., & Bangsri, A. (2020). Thai students' perceived teacher support on their reading ability: Mediating effects of self-efficacy and sense of school belonging. *International Journal of Educational Methodology*. https://eric.ed.gov/?id=EJ1254856

Curley, M., & Salmelin, B. (2018). *Framing OI2. Open Innovation 2.0*. https://link.springer.com/chapter/10.1007/978-3-319-62878-3_5

Efendi, N., Barkara, R. S., Universitas, D., Negeri, I., & Binjol, I. (2021). Study of science literacy literature in elementary school. *Jurnal Dharma PGSD, 1*(2), 57–64.

Efendi, N. (2021). *Buku Ajar Filsafat Sains*. Umsida Press. https://doi.org/10.21070/2021/978-623-6292-30-3

ELIBRARY. (n.d.). *The influence of 4C (constructive, critical, creativity, collaborative) learning model on students' learning outcomes*. https://eprints.walisongo.ac.id/id/eprint/13757

Farina, D. (2019). *The effectiveness of the learning cycle 7E model in physics learning to reduce physics concept errors of class X students*. http://repository.radenintan.ac.id/7564/1/Skripsi Full.pdf

Fauth, B., Decristan, J., Decker, A. T., Büttner, G., Hardy, I., Klieme, E., & Kunter, M. (2019). The effects of teacher competence on student outcomes in elementary science education: The mediating role of teaching quality. *Teaching and TeacherEducation, 86*, 102882. https://www.sciencedirect.com/science/article/pii/S0742051X18311296

Fenanlampir, A., Batlolona, J. R., & Imelda, I. (2019). The struggle of Indonesian students in the context of TIMSS and PISA has not ended. *International Journal of Civil Engineering and Technology, 10*, 393–406. https://www.academia.edu/download/58581048/IJCIET_10_02_042.pdf

Firdaus, H., & Purnama, S. K. (2018). The development model of badminton base technique training based of audio visual media for the beginner athlete. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan, 8*, 210–214.

OECD. (2019). *OECD skills outlook 2019: Thriving in a digital world*. OECD. https://www.voced.edu.au/content/ngv:82995

Gardner, H. (1999). *Howard Gardner, multiple intelligences, and education*. https://infed.org/howard-gardner-multiple-intelligences-and-education/

Han, S., Rosli, R., Capraro, M. M., & Capraro, R. M. (2016). The effect of science, technology, engineering and mathematics (STEM) project based learning (PBL) on students' achievement in four mathematics topics. *Journal of Turkish Science Education (TUSED)*. http://tused.org/index.php/tused/article/view/554

Handayani, N. P. S. (2020). *Development of critical thinking ability instruments and science literacy of grade V elementary school students.* https://repo.undiksha.ac.id/4353/

Heksa, S. P. A. (2021). *Sainpreneur-based science extracurriculars.* https://books.google.com/books?hl=en%5C&lr=%5C&id=vxMxEAAAQBAJ%5C&oi=fnd%5C&pg=PP1%5C&dq=ekstrakurikuler+ipa+berbasis+sainpreneur%5C&ots=whRJsJ3PFK%5C&sig=dObCo507PWY13fdjwg7jbQIXA7I

Henukh, A., Simbolon, M., Astra, I. M., & Rosdianto, H. (2021). Analysis of students' science literacy ability on heat concept. *JIPF: Jurnal Ilmu Pendidikan Fisika, 6*, 178–184. https://www.researchgate.net/profile/Haris-Rosdianto/publication/352636676_Analysis_of_Students'_Science_Literacy_Ability_on_Heat_Concept/links/60d16fa892851ca3acbafcd4/Analysis-of-Students-Science-Literacy-Ability-on-Heat-Concept.pdf

Indrawati, F. A., & Wardono, W. (2019). Effect of self efficacy on mathematical literacy ability and 4C ability formation. PRISMA. *Proceedings of the National Mathematics Seminar*.

https://journal.unnes.ac.id/sju/index.php/prisma/article/view/29307

Johnson, A. M. (2021). *Learning science: The value of crafting engagement in science environments*. https://journals.sagepub.com/doi/full/10.1177/00943061211021084v

Kartianom, K., & Retnawati, H. (2018). Why are their mathematical learning achievements different? Re-analysis TIMSS 2015 data in Indonesia, Japan and Turkey. *International Journal on New Trends in Education and Their Implications*. http://www.ijonte.org/FileUpload/ks63207/File/ijonte_2018.2_complete.pdf#page=40

Kembara, M. D., Hanny, R., Gantina, N., Kusumawati, I., Budimansyah, D., Sunarsi, D., & Khoiri, A. (2020). Scientific literacy profile of student teachers on science for all context. *Solid State Technology, 63*(6), 5844–5856. https://www.researchgate.net/profile/Ahmad-Khoiri/publication/345760061_Scientific_Literacy_Profile_Of_Student_Teachers_On_Science_For_All_Context/links/5face226a6fdcc9389ab3e35/Scientific-Literacy-Profile-Of-Student-Teachers-On-Science-For-All-Context.pd

Kristyaningrum, D. H., & Winarto, W. (2020). Elementary school teachers' perceptions of science learning practices. *Journal of Basic Education*. https://elearning.umpwr.ac.id/ejournal2/index.php/JPD/article/view/6558

Linacre, J. M. (2006). Data variance explained by Rasch measures. *Rasch Measurement Transactions*. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.693.9690%5C&rep=rep1%5C&type=pdf

Linacre, J. M. (2008). Variance in data explained by Rasch measures. *Rasch Measurement Transaction, 22*, 1164.

Mahmut, Ö. (2020). What does PISA tell us about performance of education systems? *Bartın University Journal of Faculty of Education*. https://dergipark.org.tr/en/pub/buefad/issue/53032/697153

Meuthia, H., & Ahmad, R. (2021). Analisis Aspek-Aspek Literasi Sains Pada Buku Kurikulum 2013 Tema 1 Kelas IV SD. *JEMS: Jurnal Edukasi Matematika dan Sains, 9*, 86–98. http://e-journal.unipma.ac.id/index.php/JEMS/article/view/8667

Misbach, I. H., & Sumintono, B. (2014). *Pengembangan dan validasi instrumen "persepsi siswa tehadap karakter moral guru" di Indonesia dengan model Rasch*. http://eprints.um.edu.my/10016/

Moodley, K., & Gaigher, E. (2019). Teaching electric circuits: Teachers' perceptions and learners' misconceptions. *Research in Science Education*. https://link.springer.com/article/10.1007/s11165-017-9615-5

Mufit, F., Hanum, S. A., & Fadhilah, A. (2020). Preliminary research in the development of physics teaching materials that integrate new literacy and disaster literacy. *Journal of Physics: Conference Series*. https://iopscience.iop.org/article/10.1088/1742-6596/1481/1/012041/meta

Nadelson, L. S., Heddy, B. C., Jones, S., Taasoobshirazi, G., & Johnson, M. (2018). Conceptual change in science teaching and learning: Introducing the dynamic model of conceptual change. *International Journal of Educational Psychology*. https://hipatiapress.com/hpjournals/index.php/ijep/article/view/3349

Nugrahanto, S., & Zuchdi, D. (2019). Indonesia PISA result and impact on the reading learning program in Indonesia. *International Conference on Interdisciplinary Language, Literature and Education (ICILLE 2018)* (pp. 373–377).

Nuril, L. F. (2020). *The effect of electronic portfolio assessment on improving the science literacy of class XI students in biology subjects*. http://repository.radenintan.ac.id/id/eprint/12599

Pahrudin, M. P. (2018). *Korespondensi: The analysis of pre-service physics teachers in scientific literacy: Focus on the competence and knowledge aspects*.

http://repository.radenintan.ac.id/10568/1/Korespondensi Seluruh ke JPII-Q3 %28Agus Pahrudin%29.pdf

Pahrudin, M. P. (2019). The analysis of pre-service physics teachers in scientific literacy: Focus on the competence and knowledge aspects. *Jurnal Pendidikan IPA Indonesia*. http://repository.radenintan.ac.id/10145/1/THE ANALYSIS OF PRE-SERVICE PHYSICS TEACHERS IN SCIENTIFIC LITERACY.pdf

Rohmawati, E., Widodo, W., & Agustini, R. (2018). Building students' science literacy skills through learning with a socio-scientific issues context assisted by weblog media. *Science Education Research.* https://journal.unesa.ac.id/index.php/jppipa/article/view/3146

Rusyadi, A. (2021). Guided inquiry-based science learning. *Proceedings of the Master of Science Education*. http://jbse.ulm.ac.id/index.php/PMPIPA/article/view/25

Salas-Velasco, M., Moreno-Herrero, D., & Sánchez-Campillo, J. (2021). Teaching financial education in schools and students' financial literacy: A cross-country analysis with PISA data. *International Journal of Finance & Economics*. https://onlinelibrary.wiley.com/doi/abs/10.1002/ijfe.2005

Santyasa, I. W., Rapi, N. K., & Sara, I. (2020). Project based learning and academic procrastination of students in learning physics. *International Journal of Instruction*. https://eric.ed.gov/?id=EJ1239269

Sari, R. K., Hidayat, C., & Bangapadang, S. (2021). Analysis of the role of universities in supporting sustainable growth: The case in Indonesia. *IOP Conference Series: Earth and Environmental Science*. https://iopscience.iop.org/article/10.1088/1755-1315/729/1/012093/meta

Sharon, A. J., & Baram-Tsabari, A. (2020). Can science literacy help individuals identify misinformation in everyday life? *Science Education*. https://onlinelibrary.wiley.com/doi/abs/10.1002/sce.21581

She, H. C., Stacey, K., & Schmidt, W. H. (2018). Science and mathematics literacy: PISA for better school education. *International Journal of Science and Mathematics Education*. https://link.springer.com/article/10.1007/s10763-018-9911-1

Snow, N. (2013). The meaning of science in the copyright clause. https://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/byulr2013%5C&section=13

Sumintono, B., & Widhiarso, W. (2009). Rasch model of modern measurement methods in the social sciences. Trim Komunikata Publishing.

Sumintono, B., & Widhiarso, W. (2014). *Application of rasch model for social sciences research* (revised edition). http://eprints.um.edu.my/11413/

Supena, I., Darmuki, A., & Hariyadi, A. (2021). The influence of 4C (constructive, critical, creativity, collaborative) learning model on students' learning outcomes. *International Journal of Instruction, 14*, 873–892. https://eric.ed.gov/?id=EJ1304598

Susiani, K., Suastra, I. W., & Arnyana, I. B. P. (2022). Study of improving the quality of learning in an effort to improve the quality of elementary school education. *Jurnal EDUCATIO: Jurnal Pendidikan Indonesia, 8*(1), 37–44. http://www.jurnal.iicet.org/index.php/j-edu/article/view/1786

Tze, V., & Li, J. C. H. (2021). Should science be a subject that we should keep our mind cool? A systemic evaluation using TIMSS 2015 data. *Motivation and Emotion*. https://link.springer.com/article/10.1007/s11031-021-09880-8

Warpala, I. W. S. (2006). *The influence of different learning approaches and cooperative learning strategies on understanding and critical thinking skills* [Unpublished]. State University of Malang.

Wen, C. T., Liu, C. C., Chang, H. Y., Chang, C. J., Chang, M. H., Chiang, S. H. F., Yang, C. W., & Hwang, F. K. (2020). Students' guided inquiry with simulation and its relation to school science achievement and scientific literacy. *Computers & Education, 149*, 103830.

https://www.sciencedirect.com/science/article/pii/S0360131520300324

Widiyatmoko, A., & Shimizu, K. (2018a). Literature review of factors contributing to students' misconceptions in light and optical instruments. *International Journal of Environmental and Science Education, 13*(10), 853–863. https://www.researchgate.net/profile/Arif-Widiyatmoko/publication/330102664_Literature_Review_of_Factors_Contributing_to_Students'_Misconceptions_in_Light_and_Optical_Instruments/links/5c2d93caa6fdccfc70791127/Literature-Review-of-Factors-Contributing-to-

Widiyatmoko, A., & Shimizu, K. (2018b). Factors contributing to students' misconceptions in light and optical instruments in Indonesia a literature review. *Proceedings of the Annual Meeting of the Japan Society for Science Education 42*. https://www.jstage.jst.go.jp/article/jssep/42/0/42_393/_article/-char/ja/

Wijayani, N. H. (2020). The mathematical literacy process of students in solving PISA problems in content space and shape. http://repositori.unsil.ac.id/id/eprint/3631

Yuliati, Y. (2017). Science literacy in science learning. *Pendas Horizon Journal, 3*(2), 21–28. https://doi.org/10.31949/jcp.v3i2.592