JPPIPA 9(10) (2023)

Jurnal Penelitian Pendidikan IPA

Journal of Research in Science Education

http://jppipa.unram.ac.id/index.php/jppipa/index

Science Literacy Assessment Instrument for Additives and Addictive Substances: Development, Validation, and Rasch Model Analysis

Syarif Faisal[1*], Ani Rusilowati[2], Endang Susilaningsih[2]

[1] Doctoral Program of Science Education, Universitas Negeri Semarang, Semarang, Indonesia.
[2] Natural Science Education, Universitas Negeri Semarang, Semarang, Indonesia.

Received: June 20, 2023
Revised: August 10, 2023
Accepted: October 25, 2023
Published: October 31, 2023

Corresponding Author:
Syarif Faisal
faisalidrus0@students.unnes.ac.id

DOI: 10.29303/jppipa.v9i10.4376

Abstract: Science literacy is one of the areas of basic literacy that PISA focuses on. Science literacy is essence of learning in the 21st century. The purpose of this study is to develop and validate instruments to assess science literacy in the area of additives and addictive substances. The researcher refers to the modified PISA framework and Bloom's Revised Taxonomy. The developed instrument was tested for reliability and validity using the Aiken V method and the Rasch model. The research respondents consisted of 47 students from two junior high schools: a public school and a private school that has been accredited A. The results of the analysis showed that the instrument has validity. The results of the analysis showed that of the 25 items given there were 24 valid items, the resulting Cronbach's alpha value of 0.86 showed very good criteria, it can be concluded that the instrument has high validity and good reliability. This instrument is expected to be used to measure the quality of education and science literacy of students in Indonesia.

Keywords: Bloom's revised taxonomy; Modified framework; PISA; Rasch model; Science literacy assessment instrument

Introduction

The quality of education reflects a country's progress. The difference between developed and developing countries can be seen in the development of each sector, especially the economy and economic growth (Gani et al., 2018). OECD (Organization for Economic Cooperation and Development) is an organization that focuses on economic cooperation and development. Through its PISA program, the organization helps countries prepare human resources with the skills expected in the international marketplace. PISA measures the quality of human resources through basic skills such as reading, mathematics, and science (Pratiwi, 2019). Thus, the quality of education through measured literacy will illustrate how much progress a country has made.

Science literacy is one of the areas of basic literacy that PISA focuses on. Science literacy is essence of learning in the 21st century. In addition, scientific literacy is also very important to solve various problems related to ethics, morals and global issues due to rapid changes in the field of science and technology (Jamaluddin et al., 2019). According to Nurcahyani in (Mujahidin et al., 2023) the basic abilities of students that need to be developed and mastered are scientific literacy. There are three inseparable components in science: attitude, process, and product. Assessment is needed to measure what has been achieved and to compare the quality of human resources with other countries. According to Rusilowati (2018), based on PISA data from 2000 to 2015, the scientific literacy of Indonesian children has always been below average. Therefore, efforts should be made to improve the ability of Indonesian human resources. Therefore, assessment standards are the main key to determine the ability of science literacy. According to Suparya et al. (2022), there are eight factors that cause the low learning outcomes of Indonesian students,

How to Cite:
Faisal, S., Rusilowati, A., & Susilaningsih, E. (2023). Science Literacy Assessment Instrument for Additives and Addictive Substances: Development, Validation, and Rasch Model Analysis. *Jurnal Penelitian Pendidikan IPA*, *9*(10), 7826-7836. https://doi.org/10.29303/jppipa.v9i10.4376

namely: student textbooks, student misconceptions, decontextualized learning, low reading skills, learning environment and climate, school infrastructure, human resources, and school management. Decontextualized learning and low reading ability should be the main concerns in improving science literacy. Learning that is not contextualized does not provide students with authentic tasks. Students need to be presented with certain concepts in a complex environment so that they can determine how and where those concepts can be implemented (Lotulung Mareike et al., 2017). Traditional pedagogical approaches of memorizing formulas and concepts are still found in science classrooms today (Desnita et al., 2022). The low reading proficiency of Indonesian students has an impact on student performance. Facts show that the reading ability of Indonesian people is still low and directly proportional to their cognitive ability (Kohar, 2022). This is confirmed by Indonesia's below-average PISA science literacy test results because contextual information is an assessment aspect of science literacy skills. Assessment or evaluation forms the foundation of improving education quality. To gauge student aptitude, teachers can utilize evaluations that reveal their fundamental capacity to read contextual information.

One of the efforts that can be made to improve science literacy and contextualized learning is through assessment instruments based on the PISA science literacy framework and adapted to a country's curriculum. Assessment instruments based on the PISA framework will be able to predict the science literacy skills of students in Indonesia at an early stage. According to Cansız et al. (2019), there needs to be a balance between the science curriculum and the PISA science literacy framework. The PISA science literacy framework consists of four aspects: context, knowledge, competence, and attitude. The balance between the PISA science literacy dimensions and the curriculum needs to be improved in order to help students become internationally science literate.

In order to assess students' scientific literacy, a test can be administered in the form of questions that follow the PISA framework, which includes the identification of context, competence, and knowledge (Heuston, 2022). Context is the three parts that relate to human life: personal, local/national, and global. Competencies include the ability to explain scientific phenomena, design and evaluate scientific investigations, and interpret data. Knowledge refers to content, process, and epistemic knowledge (She et al., 2018).

To determine the depth of knowledge of the science literacy test used, it is necessary to divide the level of difficulty of all the questions created. The difficulty level is divided into low, medium, and high (Heuston, 2022). The PISA framework for creating science literacy test questions is shown in Figure 1 below.

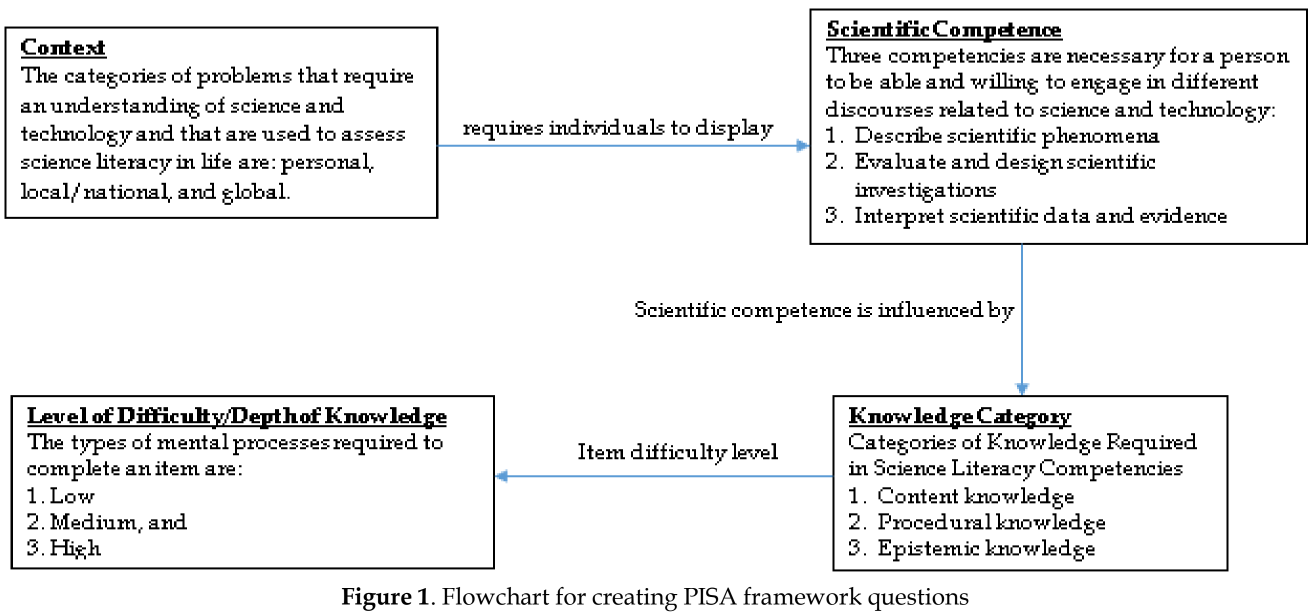


Figure 1. Flowchart for creating PISA framework questions

The PISA framework needs to be aligned with the assessment framework applied to a country's curriculum. Arlianty et al. (2018) stated that several education observers explained the tendency of learning in Indonesia to only explain knowledge, laws, and facts, but not to relate the material to everyday life. The Indonesian government, through the Education Quality Assurance Agency (LPMP), emphasizes students' critical thinking through learning. The Indonesian government also decided to increase the analysis test by 10 percent each year. In practice, teachers are advised to ask questions to improve higher order thinking skills

(HOTS). To achieve this higher order thinking, there must be a framework that can be used. The framework that can be used is Bloom's taxonomy or Bloom's revised taxonomy which divides the levels of thinking from low order thinking skill to higher order thinking skill with six dimensions from C1 to C6. The use of books in Indonesia applies the use of Bloom's revised or Anderson's taxonomy thinking dimensions since 2013. These levels of thinking are remembering, understanding, applying, analyzing, evaluating, and creating (Febrina et al., 2019). The Indonesian government has taken the right step by improving students' critical thinking skills towards higher order thinking skills through the use of the revised Bloom's taxonomy framework. This step needs to be balanced with the daily use of the PISA framework so that there is a balance between the adaptation of the PISA framework and the revised Bloom's Taxonomy framework as shown in Figure 2.

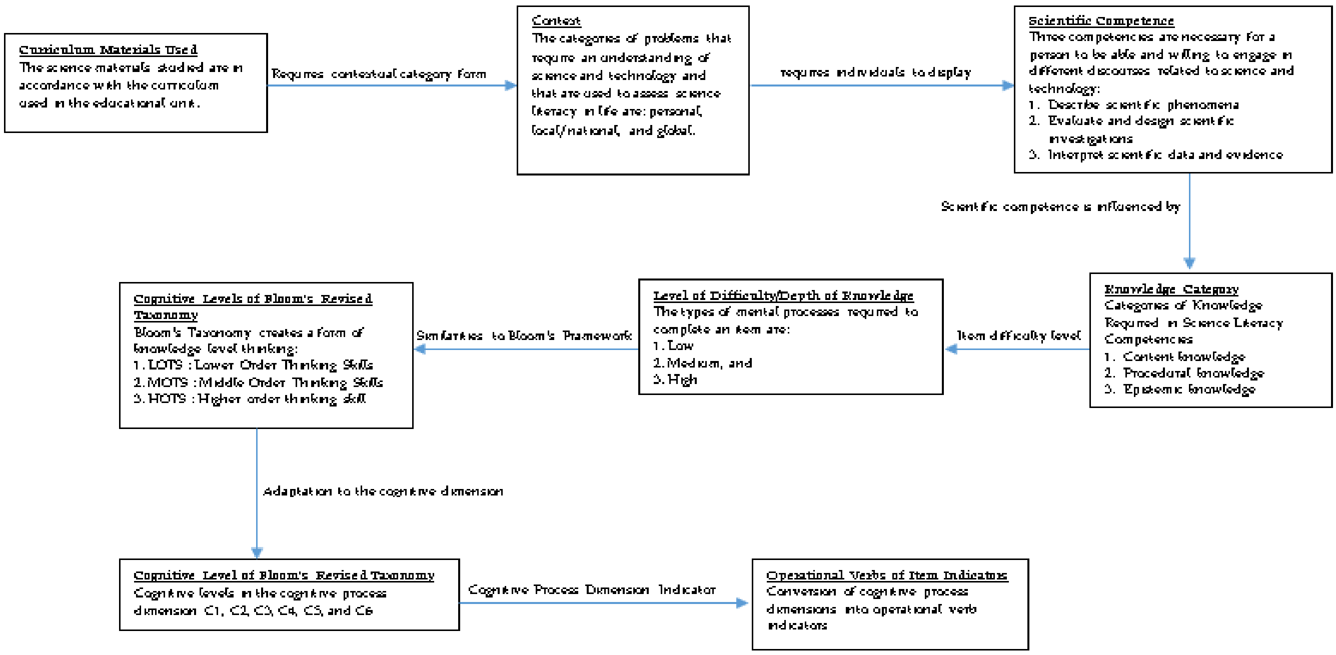


Figure 2. Flowchart of question modification based on PISA framework and Bloom's revised taxonomy framework

Many studies have been conducted on science literacy, such as research on the science literacy profile of science teachers in junior high schools (Jamaluddin et al., 2019), implementation of blended learning using STEM to improve science literacy (Lestari et al., 2021), development of an assessment sheet to measure students' science literacy level (Zahara et al., 2022), making stem-based e-modules to improve science literacy (Oktarina et al., 2023), worksheet application to improve students' science literacy skills (Sahnaz & Kuswandi, 2023). In addition, instrument development using the Rasch model has also been commonly carried out in the field of literacy (Ad'hiya & Laksono, 2018; Bahri et al., 2021; El Masri & Andrich, 2020; Koerber & Osterhaus, 2019; Sadhu & Laksono, 2018; Setyorini et al., 2021; Susongko et al., 2021; Wati et al., 2018; Woudstra et al., 2019). However, no research exists on the modification of the PISA framework using the revised Bloom's taxonomy framework for the development of science literacy assessment tools. The prepared questions will then be tested for reliability and validity, so that questions related to the modified framework and the modified framework can be used as instruments for predicting the early science literacy of students in Indonesia. It is expected that the questions referring to the modified framework can be used as an instrument in predicting students' science literacy skills early.

This study aims to develop a science literacy questionnaire based on the modified framework presented in Figure 2. The prepared questions are subsequently evaluated for reliability and validity through the application of the Rasch model.

Method

The research methodology employed is that of development research. As per Restu, development research yields practical products which are a summarized outcome of research and development (Oktarina et al., 2023).

The following are the steps taken in this research: 1) Needs analysis of science literacy outcomes in Indonesia; 2) Design a modified framework from the PISA framework and Bloom's revised taxonomy; 3)

Develop evaluation instruments from the modified framework; 4) The instruments developed are assessed for feasibility and validity using Aiken V and Rasch Model analysis.

The prepared questions are then tested for reliability and validity, so that questions related to the modified framework and modified framework can be used as instruments in predicting early science literacy of students in Indonesia. The respondents of this study amounted to 47 students from two junior high schools selected using purposive sampling method. The consideration is that both schools are public schools SMPN 71 Jakarta and private SMP Mazaya Assunnah, which represent the form of schools in Indonesia and both schools have been accredited A. Limited time and resources are also a consideration for this method, which aims to make the research faster and more efficient to carry out. Respondents were not coerced and did not receive any consequences for participating in this study. The form of the graph is presented in Table 1, with restrictions on the presentation of the PISA scientific literacy and Bloom's revised taxonomy levels and the number of questions. The purpose of the presentation is to show the relationship between the two frameworks in a simple form.

Table 1. Simple Representation of Modified PISA and Bloom's Taxonomy Science Literacy Grids

PISA Scientific Competence	Revised Bloom's Taxonomy Level	Question Number
Interpret data and evidence scientifically	LOTS	5, 6, 21
	MOTS	1, 16, 20
Explain scientific phenomena	LOTS	3, 7, 11, 15, 22, 23, 24, 25
	MOTS	4, 9, 12, 18
	HOTS	2, 8, 10, 17
Evaluate and design scientific studies	HOTS	13, 14, 19

In Table 1, a simple representation shows that there are 6 questions for the competency to interpret data and evidence scientifically. While to explain scientific phenomena there are 12 questions. And to evaluate and design scientific investigations, there are 3 questions. From the revised Bloom taxonomy level, there are 11 questions for LOTS, 7 questions for MOTS, and 7 questions for HOTS.

9. Pada bulan Januari 1985, institut kami merawat seorang pria berusia 34 tahun bernama S.C. yang menderita pembengkakan berulang pada bibir atas dan gusi, yang terkait dengan adanya lidah yang retak. Kejadian pertama pembengkakan terjadi ketika pasien berusia 10 tahun dan disebabkan oleh konsumsi keju berbumbu. Sejak April 1980, kejadian tersebut menjadi lebih sering dan terjadi setelah pasien mengonsumsi makanan tertentu, termasuk anggur, sau-sage, dan makanan pedas. Pembengkakan selalu mereda secara spontan dalam waktu dua minggu. Pada tahun 1981, pasien melakukan tes tusuk kulit untuk dermatophagoides pteronissimus dan dermatophagoides farinae, yang positif, sehingga pasien menjalani imunoterapi spesifik. Namun, setelah dua tahun pengobatan, terapi dihentikan karena tidak memberikan manfaat.

Saudara perempuan pasien memiliki lidah yang berlipat, namun tidak ada riwayat penyakit alergi dalam keluarga. Berdasarkan riwayat pasien, pemeriksaan fisik, dan histologi lesi, pasien didiagnosis menderita MRS. Tes tusuk kulit positif terhadap sereal menunjukkan kemungkinan adanya mekanisme alergi pada kelainan ini. Oleh karena itu, pasien menjalani program diet terbatas selama tiga minggu, yang menyebabkan pembengkakan berangsur membaik. Tersangka penyebab pembengkakan adalah anggur dan bahan tambahan makanan, yang diuji dengan uji tantang. Uji tantang tersebut positif, sehingga terapi dilanjutkan dengan ketotifen fumarat dan kromoglikat. Meskipun pasien membaik, kejadian sindrom tidak sepenuhnya teratasi.

Pada bulan Juni 1986, pasien diuji ulang untuk intoleransi makanan dengan diet eliminasi dan uji tantang buta ganda untuk bahan tambahan makanan. Gejala-gejala yang dialami pasien disebabkan oleh sodium benzoate dan tartrazine. Aditif tersebut kemudian dihilangkan dari makanan pasien, dan penyakitnya semakin membaik hingga sembuh total dan telah berlangsung selama lebih dari satu tahun.

Sumber : (Pachor et al., 1989)

Penyebab utama terjadinya pembengkakan pada pasien di sebabkan oleh ….

a. anggur dan ketotifen fumarat
b. makanan pedas dan sodium benzoate
c. kromoglikat dan ketotifen fumarat
d. tartrazine dan sodium benzoate

Figure 3. Sample science literacy questions from the developed instrument

The created instrument follows the framework in Figure 2. Examples of questions from the scientific literacy instrument on additive material are shown in Figure 3.

To ask the question, one must go through the framework in Figure 2 by identifying the material, namely additives, which are material at the junior high school level. The context that wants to be raised on the material is adapted to the data taken contextually. The context taken is personal. The competence to be achieved is the explanation of scientific phenomena. The knowledge to be revealed is in the form of content, with a moderate level of depth of knowledge and in accordance with the MOTS in Bloom's revised taxonomy. The dimension of Bloom's revised taxonomy is C3 with the indicator that students can determine the impact of the use of preservatives on food.

Aiken V analysis and Rasch model were used in this study. Aiken V was used to determine the content validity of the instrument. The instrument was validated by six validators consisting of four lecturers and two teachers. The questionnaire given to the validators used a five-point Likert scale, namely 1-5 (Marin et al., 1987). Many instrument measurements have been carried out using the Rasch model because it can measure latent human cognitive abilities (Ariffin et al., 2010; Mulyanti et al., 2022). The Rasch model can see the relationship between respondents and items simultaneously. The scores obtained from the Rasch model are logit scores, not raw scores. The logit value reflects the probability of an item being selected by a group of respondents. The purpose of using this model is to anticipate the ordinal form of raw scores from Likert scales that do not have equal intervals between scores (Muntazhimah et al., 2020). This instrument uses dichotomous data processing in the Rasch model because the question is a multiple choice question that requires respondents to select one of the correct answers. The number of dichotomous scores becomes a statistic to estimate the respondent's ability and the difficulty of the question items on a logit scale (Tesio et al., 2023). The summary statistic output is used as the data analysis output in Figure 5 to determine the reliability of the instrument. While the validity of the instrument uses the output of item undimensionality and item fit order.

Result and Discussion

Aiken V

The results obtained from the distribution of Likert scale questionnaires to validators were analyzed using the Aiken V index. The Aiken V formula for analyzing the content validity of this questionnaire is as follows:

$$V = \frac{\sum s}{n(c-1)} \qquad (1)$$

Description:
V : rater agreement index regarding item validity
S : r- Io
R : the rating by an expert
Io : the lowest possible validity rating
c : number of levels of the rating scale used
n : number of judges

Question Number	Expert Validator						s1	s2	s3	s4	s5	s6	SS	v	Category
	I	II	III	IV	V	VI									
1	5	5	5	5	4	4	4	4	4	4	3	3	22	0,92	Valid
2	5	5	5	5	5	4	4	4	4	4	4	3	23	0,96	Valid
3	5	5	5	5	4	4	4	4	4	4	3	3	22	0,92	Valid
4	5	5	5	5	4	4	4	4	4	4	3	3	22	0,92	Valid
5	5	5	4	5	5	5	4	4	3	4	4	4	23	0,96	Valid
6	5	5	5	5	4	4	4	4	4	4	3	3	22	0,92	Valid
7	5	5	4	5	5	4	4	4	3	4	4	3	22	0,92	Valid
8	5	5	5	5	5	4	4	4	4	4	4	3	23	0,96	Valid
9	5	5	5	5	4	4	4	4	4	4	3	3	22	0,92	Valid
10	5	5	5	5	4	5	4	4	4	4	3	4	23	0,96	Valid
11	5	5	5	5	5	5	4	4	4	4	4	4	24	1,00	Valid
12	5	5	5	5	4	5	4	4	4	4	3	4	23	0,96	Valid
13	5	5	5	4	4	4	4	4	4	3	3	3	21	0,88	Valid
14	5	5	5	5	5	4	4	4	4	4	4	3	23	0,96	Valid
15	5	5	5	4	4	5	4	4	4	3	3	4	22	0,92	Valid
16	5	5	4	5	4	5	4	4	3	4	3	4	22	0,92	Valid
17	5	5	5	5	4	4	4	4	4	4	3	3	22	0,92	Valid
18	5	5	5	5	4	5	4	4	4	4	3	4	23	0,96	Valid
19	5	5	5	5	4	4	4	4	4	4	3	3	22	0,92	Valid
20	5	5	4	4	4	5	4	4	3	3	3	4	21	0,88	Valid
21	5	5	5	5	4	4	4	4	4	4	3	3	22	0,92	Valid
22	5	5	5	5	5	5	4	4	4	4	4	4	24	1,00	Valid
23	5	5	5	5	4	5	4	4	4	4	3	4	23	0,96	Valid
24	5	5	5	5	4	5	4	4	4	4	3	4	23	0,96	Valid
25	5	5	5	4	4	5	4	4	4	3	3	4	22	0,92	Valid

Figure 4. Aiken V average score table

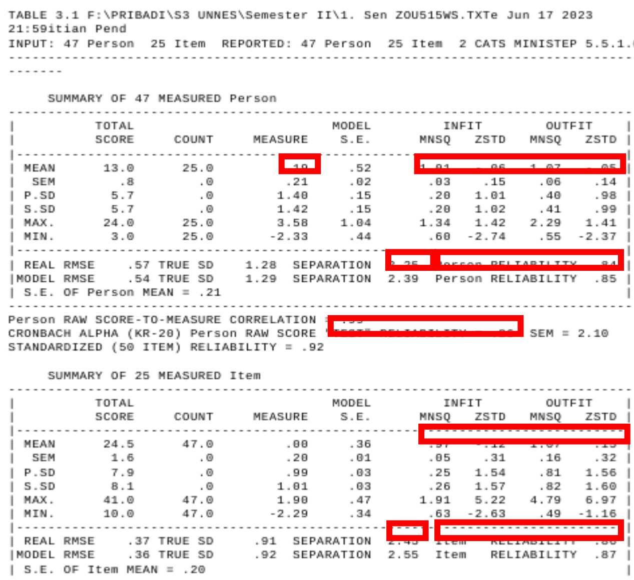
The formula is processed using the Excel application, which results in the agreement index of the validators. The question is said to be valid if the final result of the Aiken V analysis meets the threshold of the Aiken V index coefficient. From the data obtained, the final Aiken V scores for 25 questions show high validity because they exceed the 0.8 threshold of the Aiken V index. The final Aiken V score is shown in Figure 4.

Rasch Model

Rasch model analysis data were processed using Ministep 5.5.1.0 software, which is a free application from Winstep. This software can only process data from 25 question items and 75 respondents (Rusilowati, 2018). The questions compiled in this science literacy instrument consisted of 25 question items with 47 respondents. Therefore, the data obtained from the respondents can be processed using Ministep software.

Instrument Reliability

The reliability of the instrument is to see if the science literacy instrument with the PISA framework and Bloom's revised taxonomy can be used to measure the science literacy of junior high school students. To see the reliability of the instrument, the summary statistics output is used (see Figure 5).


```
TABLE 3.1 F:\PRIBADI\S3 UNNES\Semester II\1. Sen ZOU515WS.TXTe Jun 17 2023
21:59itian Pend
INPUT: 47 Person  25 Item  REPORTED: 47 Person  25 Item  2 CATS MINISTEP 5.5.1.(
-------------------------------------------------------------------------------
-------

     SUMMARY OF 47 MEASURED Person
-------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT    MEASURE     S.E.      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      13.0      25.0    [illegible]  .52     [illegible]                  |
|  SEM        .8        .0        .21      .02       .03    .15    .06    .14 |
| P.SD       5.7        .0       1.40      .15       .20   1.01    .40    .98 |
| S.SD       5.7        .0       1.42      .15       .20   1.02    .41    .99 |
| MAX.      24.0      25.0       3.58     1.04      1.34   1.42   2.29   1.41 |
| MIN.       3.0      25.0      -2.33      .44       .60  -2.74    .55  -2.37 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .57 TRUE SD    1.28  SEPARATION [illegible]                     |
|MODEL RMSE    .54 TRUE SD    1.29  SEPARATION  2.39  Person RELIABILITY  .85 |
| S.E. OF Person MEAN = .21                                                   |
-------------------------------------------------------------------------------
Person RAW SCORE-TO-MEASURE CORRELATION = [illegible]
CRONBACH ALPHA (KR-20) Person RAW SCORE [illegible] SEM = 2.10
STANDARDIZED (50 ITEM) RELIABILITY = .92

     SUMMARY OF 25 MEASURED Item
-------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT    MEASURE     S.E.      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      24.5      47.0        .00      .36     [illegible]                  |
|  SEM       1.6        .0        .20      .01       .05    .31    .16    .32 |
| P.SD       7.9        .0        .99      .03       .25   1.54    .81   1.56 |
| S.SD       8.1        .0       1.01      .03       .26   1.57    .82   1.60 |
| MAX.      41.0      47.0       1.90      .47      1.91   5.22   4.79   6.97 |
| MIN.      10.0      47.0      -2.29      .34       .63  -2.63    .49  -1.16 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .37 TRUE SD     .91  SEPARATION [illegible]                     |
|MODEL RMSE    .36 TRUE SD     .92  SEPARATION  2.55  Item   RELIABILITY  .87 |
| S.E. OF Item MEAN = .20                                                     |
-------------------------------------------------------------------------------
```

Figure 5. Summary statistics output from Ministep

Information that can be interpreted through the summary statistics output to measure reliability in terms of both respondents and items and the interaction between the two. The following summary statistics figure shows the reliability information of the instrument used.

The output in Figure 5 provides information about the quality of the question instrument and the respondents in answering the question and their interactions. It should be discussed in Figure 5 that the person measure shows the average value of the respondents seen from the mean measure score of 0.19. To measure the reliability of the instrument, Cronbach's alpha value is used, which is the interaction between respondents and questions with a value of 0.86. The person reliability value shows the consistency of respondents' answers which is 0.84. The person reliability value of 0.84 shows very good results, according to Sukmawati et al. (2023) this can be interpreted that there is student consistency in answering question items and the sensitivity of question items in measuring all categories of students. The item reliability value, which is the reliability of the question items in knowing the quality of these items, is 0.86. Similarly, we can see the infit and outfit of MNSQ and ZSTD and their separation.

Instrument Validity

Instrument validity is used to test whether the science literacy instrument, which was created based on the PISA framework and Bloom's revised taxonomy, can measure the science literacy skills of students, especially at the junior high school level, according to the Additives and Addictive Substances material created in this instrument. This is done to determine whether the framework used can explain the interaction between respondents and question items.

```
TABLE 23.0 F:\PRIBADI\S3 UNNES\Semester II\1. Se ZOU515WS.TXTm Jun 17 2023
21:59litian Pend
INPUT: 47 Person  25 Item  REPORTED: 47 Person  25 Item  2 CATS MINISTEP 5.5.1.0
--------------------------------------------------------------------------------
--------

     Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information
units
                                          Eigenvalue   Observed   Expected
Total raw variance in observations     =    37.8835 100.0%          100.0%
  Raw variance explained by measures   =    12.8835  34.0%           33.1%
    Raw variance explained by persons  =     7.6000  20.1%           19.5%
    Raw Variance explained by items    =     5.2835  13.9%           13.6%
  Raw unexplained variance (total)     =    25.0000  66.0% 100.0%    66.9%
    Unexplned variance in 1st contrast =     2.8791   7.6%  11.5%
    Unexplned variance in 2nd contrast =     2.3860   6.3%   9.5%
    Unexplned variance in 3rd contrast =     2.1881   5.8%   8.8%
    Unexplned variance in 4th contrast =     2.0852   5.5%   8.3%
    Unexplned variance in 5th contrast =     1.7941   4.7%   7.2%
```

Figure 6. Item undimensionality output from ministep

Ministep shows the output of item undimensionality in Figure 6 and item fit order in Figure 7. According to Darmana et al. (2021) unidimensional is also called the construct validity of the instrument. The output shows the test of instrument items, i.e. question items that can measure what should be measured. The analysis of the validity of this science literacy instrument in the Ministep program is called a fit or misfit test, which is performed by analyzing the output of this item fit order. The following output shows information about the validity criteria of the instrument used.

```
TABLE 10.1 F:\PRIBADI\S3 UNNES\Semester II\1. Se ZOU515WS.TXTm Jun 17 2023 21:59litian Pend
INPUT: 47 Person  25 Item  REPORTED: 47 Person  25 Item  2 CATS MINISTEP 5.5.1.0
--------------------------------------------------------------------------------------------
Person: REAL SEP.: 2.25  REL.: .84 ... Item: REAL SEP.: 2.45  REL.: .86

        Item STATISTICS:  MISFIT ORDER

--------------------------------------------------------------------------------------------
|ENTRY   TOTAL  TOTAL    JMLE    MODEL|   INFIT   |  OUTFIT   |PTMEASUR-AL|EXACT  MATCH|      |
|NUMBER  SCORE  COUNT  MEASURE  S.E.  |MNSQ  ZSTD|MNSQ  ZSTD|CORR.   EXP.| OBS%  EXP%| Item |
|-------------------------------------+----------+----------+-----------+-----------+------|
|    24     26     47     -.17     .34|1.91  5.22|4.79  6.97|A-.26    .48| 46.8  72.2| i24  |
|    23     31     47     -.75     .35|1.14   .91|1.63  1.46|B .34    .45| 68.1  74.3| i23  |
|     8     24     47      .06     .34|1.39  2.56|1.25   .94|C .28    .49| 53.2  71.3| i8   |
|     2     16     47     1.01     .36|1.11   .69|1.37  1.20|D .42    .51| 70.2  75.3| i2   |
|    13     27     47     -.28     .34|1.18  1.23|1.13   .50|E .37    .48| 70.2  72.6| i13  |
|    22     20     47      .52     .34|1.12   .86|1.13   .56|F .43    .51| 74.5  72.1| i22  |
|    19     20     47      .52     .34|1.01   .14|1.09   .43|G .48    .51| 74.5  72.1| i19  |
|    12     24     47      .06     .34|1.06   .46| .93  -.18|H .47    .49| 66.0  71.3| i12  |
|    14     34     47    -1.14     .37| .96  -.15|1.01   .18|I .43    .42| 78.7  76.3| i14  |
|    18     20     47      .52     .34|1.01   .11| .98   .01|J .50    .51| 74.5  72.1| i18  |
|     4     31     47     -.75     .35| .78 -1.51|1.00   .14|K .56    .45| 85.1  74.3| i4   |
|    20     12     47     1.57     .39| .88  -.53|1.00   .14|L .56    .51| 85.1  81.2| i20  |
|     1     15     47     1.14     .37| .94  -.27| .97   .02|M .54    .51| 78.7  76.5| i1   |
|    15     23     47      .17     .34| .94  -.40| .86  -.45|l .54    .50| 72.3  71.1| i15  |
|     3     33     47    -1.01     .36| .92  -.41| .81  -.29|k .48    .43| 74.5  75.5| i3   |
|     6     41     47    -2.29     .47| .92  -.17| .53  -.40|j .39    .31| 85.1  87.2| i6   ||
|     9     28     47     -.40     .34| .82 -1.25| .92  -.13|i .56    .47| 78.7  73.0| i9   |
|    11     34     47    -1.14     .37| .89  -.60| .69  -.55|h .50    .42| 74.5  76.3| i11  |
|    16     19     47      .64     .35| .89  -.74| .86  -.45|g .57    .51| 76.6  72.8| i16  |
|    17     16     47     1.01     .36| .87  -.78| .72  -.94|f .61    .51| 74.5  75.3| i17  |
|    25     25     47     -.05     .34| .81 -1.39| .71 -1.04|e .61    .49| 76.6  71.8| i25  |
|     5     16     47     1.01     .36| .78 -1.36| .69 -1.08|d .65    .51| 78.7  75.3| i5   |
|    21     10     47     1.90     .42| .70 -1.23| .52 -1.12|c .68    .51| 89.4  84.4| i21  |
|     7     35     47    -1.27     .37| .69 -1.84| .49 -1.04|b .60    .41| 85.1  77.5| i7   |
|    10     32     47     -.88     .35| .63 -2.63| .55 -1.16|a .66    .44| 91.5  74.7| i10  |
|-------------------------------------+----------+----------+-----------+-----------+------|
| MEAN    24.5   47.0      .00     .36| .97  -.12|1.07   .15|           | 75.3  75.1|      |
| P.SD     7.9     .0      .99     .03| .25  1.54| .81  1.56|           |  9.8   3.9|      |
--------------------------------------------------------------------------------------------
```

Figure 7. Item fit order output from ministep

Instrument validity is used to test whether the science literacy instrument, which was created based on the PISA framework and Bloom's revised taxonomy, can measure the science literacy skills of students, especially at the junior high school level, according to the Additives and Addictive Substances material created in this instrument. This is done to determine whether the framework used can explain the interaction between respondents and question items. Ministep shows the output of item undimensionality in Figure 6 and item fit order in Figure 7. The output shows the test of instrument items, i.e. question items that can measure what should be measured. The analysis of the validity of this science literacy instrument in the Ministep program is called a fit or misfit test, which is performed by analyzing the output of this item fit order. The following output shows information about the validity criteria of the instrument used.

The item statistic results on the outfit values show that i24 has quite high MNSQ and Z values. This is because some high-ability students were unable to answer the test item (Susac et al., 2018). This led to a rather low differentiating power for this item (Rakkapao et al., 2016).

Content validity using Aiken V analysis with six validators, four lecturers and two teachers, yielded an average score of more than 0.8, so that all question items were valid with a high category. The lowest score was 0.8 for question items 13 and 20, while the highest score was 1.00 for question items 11 and 22. All validators indicated that all question items were valid overall. Instruments that have been declared valid by the validators are given to respondents to determine reliability and validity using the Rasch model. Reliability assesses what is judged to be a stable instrument. The results of the instrument are said to be reliable if they produce stable results (Muntazhimah et al., 2020). The most important thing that an instrument must meet is the validity and reliability test used in a study in order for it to have a trustworthy value. Using Figure 5 Summary Output on the Science Literacy Instrument, the above results are examined as follows: The person measure shows a score of 0.19, which is greater than logit 0.0. A score greater than logit indicates that the respondent's ability tends to be greater than the difficulty level of the question or that the respondent is able to answer the questions contained in the instrument. The Cronbach's alpha value in the output of Figure 5 gives a value of 0.86. This interpretation can be seen in Table 2. The score shows excellent criteria, so the instrument used is reliable.

Table 2. Interpretation of Reliability Based on Cronbach Alpha Value (Sumintono & Widhiarso, 2015)

Score	Interpretation
a > 0.8	Very high
0.7 < a ≤ 0.8	High
0.6 < a ≤ 0.7	Fair
0.5 < a ≤ 0.6	Low
a < 0.5	Very low

The item reliability and person reliability values in the output above show values of 0.86 and 0.84. The item reliability value shows criteria of Good. Likewise, the person reliability has the same criteria as the item reliability, namely good, so it can be said that the quality of the items used in the instrument is very reliable. From the Person and Item tables in Figure 5, we can see the INFIT MNSQ and OUTPIT MNSQ values, as well as the INFIT ZSTD and OUTPIT ZSTD values. The INFIT MNSQ person value is 1.01 and the item value is 0.97; the OUTPIT MNSQ person and item values are 1.07. For good, INFIT MNSQ and OUTPIT MNSQ are 1.00. Anything close to 1.00 is said to be of good quality. For INFIT ZSTD and OUTFIT ZSTD in the person table, the average values are -0.06 and -0.05. In the INFITZTSD and OUPIT ZTSD values for the item values with the average are 0.97 and 1.07. The ideal value of INFITZTSD and OUPIT ZTSD is 0.0 in the sense that the closer to the ideal value, the better the quality. From the table we can see that the person value is close to a good value, but the item value is still far away. This means that there are items with poor quality. The grouping of people and items is known by the separation value. The higher the separation score, the better the overall quality of the instrument for both respondents and items. In Figure 5, the separation score for the person value is 2.25 and the item value is 2.45.

The above review provides a conclusion on the level of reliability for science literacy instruments based on the PISA framework and Bloom's Revised Taxonomy, as detailed in Table 3.

Table 3. Processing Reliability Test Results

Croncbach Alpha	Interpretation	Item Reliability	Interpretation	Person Reliability	Interpretation	Conclusion
0.86	Very High	0.86	Good	0.84	Good	Reliable

In Table 3, the interpretation of Cronbach Alpha is "very high". The items and the respondents have shown compatibility. The consistency of the respondents' answers was "good" and the quality of the items was also "good". The conclusion obtained is that the science literacy instrument based on the PISA framework and Bloom's Revised Taxonomy on the material of additives and addictive substances is considered reliable.

The extent to which question items are able to measure what they are intended to measure, using scientific validity quality standards (Rohmah et al., 2023). For the validity test using the Rasch model in Ministep software, it can be seen in the item undimensionality output. Item undimensionality is a measure of whether the instrument is able to validly measure what it is supposed to measure. Principal component analysis is a Rasch analysis of the standardized residual variance (in eigenvalue units) (Sumintono & Widhiarso, 2015). The raw variance explained by measures value shows the validity test for item undimensionality. Based on the raw variance explained by measures value, a value of > 20% is considered to be met, > 40% is considered to be good, and > 60% is considered to be excellent.

To find out which items are problematic and inappropriate, look at the eigenvalue and the observed value in the unexplained variance 1 contrast. The eigenvalue should be less than 3. There are no problematic items and the observed value must be less than 15%, indicating that the items are fit. The results of the validity processing using Ministep software version 5.5.1.0 are shown in Table 4.

Table 4. Processing Instrument Validity Results

Raw Variance Explained by measures	Interpretation	Unexplained variance 1st contrast		Interpretation
		Eigenvalue	observed	
34.00	considered to be met	2.88	7.6	There are no problem items

To find out which items are problematic and inappropriate, look at the eigenvalue and the observed value in the unexplained variance 1 contrast. The eigenvalue should be less than 3. There are no problematic items, and the observed value must be less than 15%, indicating that the items are appropriate. The results of validity processing using Ministep software version 5.5.1.0 are shown in Table 4. Item fit can account for normal item functioning through the correlation values of the outfit mean squares, outfit z-standards, and point measures (Bond, 2015). The criteria used to test item fit are presented in Table 5.

If three criteria in Table 5 are met, the item can be said to be "appropriate" and it is certain that the quality of the item is good and can be used. If only two criteria are met, or if only one criterion is met, then the item can be retained so that it can be classified as and can be used. If all three criteria are not met, then the item is "inappropriate" and must be corrected or replaced. Table 6 is used to determine the suitability of the item criteria.

Table 5. Criteria for Judging the Suitability of Question Items

Criteria	Score
Outfit mean square (MNSQ)	0.5 < MNSQ < 1.5
Outfit Z-standart (ZSTD)	-2.0 < ZSTD < +2.0
Point Measure Correlation	0.4 < PT Measure Corr < 0.85

Table 6. Processing of Item Fit Order Results

Question Item	Outfit		PT Measure Corr	Status achieved	Interpretation
	MNSQ	ZFTD			
1	0.97	0.02	0.54	3 Criteria	appropriate
2	1.37	1.20	0.42	3 Criteria	appropriate
3	0.81	- 0.29	0.48	3 Criteria	appropriate
4	1.00	0.14	0.56	3 Criteria	appropriate
5	0.69	-1.08	0.65	3 Criteria	appropriate
6	0.53	-0.40	0.39	3 Criteria	appropriate
7	0.49	-1.04	0.60	2 Criteria	appropriate
8	1.25	0.94	0.28	2 Criteria	appropriate
9	0.92	-0.13	0.56	3 Criteria	appropriate
10	0.55	-1.16	0.66	3 Criteria	appropriate
11	0.69	-0.55	0.50	3 Criteria	appropriate
12	0.93	-0.18	0.47	3 Criteria	appropriate
13	1.13	0.50	0.37	2 Criteria	appropriate
14	1.01	0.18	0.43	3 Criteria	appropriate
15	0.86	-0.45	1.54	2 Kriteria	appropriate
16	0.86	-0.45	0.57	3 Criteria	appropriate
17	0.72	-0.94	0.61	3 Criteria	appropriate
18	0.98	0.01	0.50	3 Criteria	appropriate
19	1.09	0.43	0.48	3 Criteria	appropriate
20	1.00	0.14	0.56	3 Criteria	appropriate
21	0.52	-1.12	0.68	3 Criteria	appropriate
22	1.13	0.56	0.43	3 Criteria	appropriate
23	1.63	1.46	0.34	1 Criteria	appropriate
24	4.79	6.97	-0.26	-	inappropriate
25	0.71	-1.04	0.61	3 Criteria	appropriate

Table 6 provides information about the validation of the items in the instrument. Of the 25 items, there are 19 items that meet three criteria, four items that meet two criteria, one item that meets one criterion, and one item that does not meet any criteria. Only 24 questions can be validly used as a science literacy instrument based on the PISA framework and Bloom's Revised Taxonomy on Additives and Addictive Substances. One of the 24 questions cannot be retained and must be discarded or replaced. The interpretation results show that there are no problematic questions, but the item fit order results show that there is one question that does not fit. This difference is to answer the results of raw variance explained by measures 34% on the criteria met, so there may be questions that are not related at all or have different variance factors. It can be seen that the MNSQ, ZTSD, and PT Measure Corr values for question number 24 are well above the predetermined threshold criteria. This confirms that item #24 is an item that measures other variances and therefore cannot be retained or replaced. The framework also classifies the items based on the depth of knowledge into easy, medium, and difficult in the PISA framework, which corresponds to LOTS, MOTS, and HOTS in Bloom's taxonomy. We can see the Rasch model of the depth of the question item instrument using Figure 8 on the Wright map.

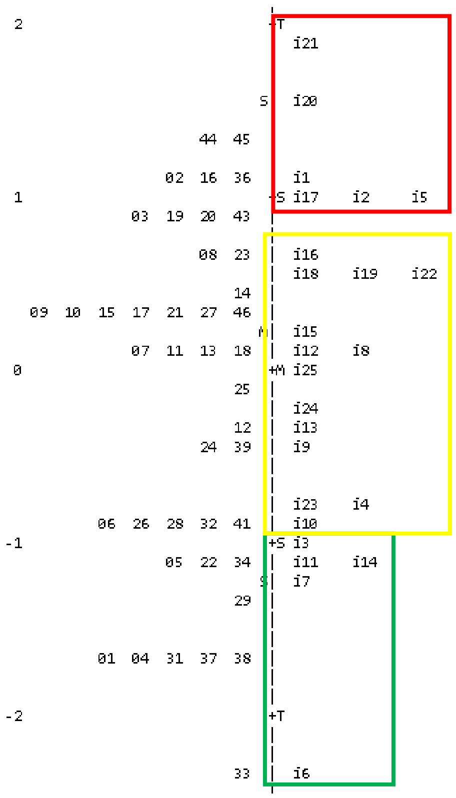


Figure 8. Output Wright Map

The Wright Map output provides information about the depth of the question items. The depth of the question is divided into three categories: easy, medium, and difficult. See Table 7 for more details.

Table 7. The Depth of the Question Item

Question Item Depth	Question Item Number
easy	3, 7, 11, 14 and 16
medium	4,8, 9, 10, 12, 13, 15, 16, 18, 19, 22, 23, and 25
difficult	1, 2, 5, 17, 20, and 21

The science literacy items in Table 7 are divided into five easy, 13 medium, and six difficult items. The science literacy item instruments based on the PISA framework and Bloom's Revised Taxonomy on Additives and Addictive Substances material have an even level of depth, namely easy, medium, and difficult.

Conclusion

The science literacy assessment instrument, which was developed based on the PISA framework and Bloom's Revised Taxonomy on Additives and Addictive Substances material, showed a Cronbach's alpha value of 0.86, indicating "excellent" criteria and high item and person reliability values. Based on the results of the Rasch model analysis, there is one item that does not fit the construct, so the questions that can be used are 24 out of the 25 compiled question items. The instrument can be used because it is considered valid and reliable.

Acknowledgment

The author would like to thank all those who have contributed to the writing of this article. I would especially like to thank Prof. Dr. Ani Rusilowati, M.Pd. and Dr. Endang Susilaningsih M.S who have provided direction in the development of science literacy assessment instruments and writing this article.

Author Contributions

This research contributes to the development of science literacy instruments using modified PISA framework and Bloom's Revised Taxonomy that can be used to measure students' science literacy skills. The roles of the authors in this research are divided into executor and advisor in this research.

Funding

This research received no external funding.

Conflicts of Interest

The author declares no conflict of interest.

References

Ad'hiya, E., & Laksono, E. W. (2018). Development and validation of an integrated assessment instrument to assess students' analytical thinking skills in chemical literacy. *International Journal of Instruction*, *11*(4), 241–256. https://doi.org/10.12973/iji.2018.11416a

Ariffin, S. R., Omar, B., Isa, A., & Sharif, S. (2010). Validity and reliability Multiple Intelligent item using Rasch measurement model. *Procedia - Social and Behavioral Sciences*, *9*, 729–733. https://doi.org/10.1016/j.sbspro.2010.12.225

Arlianty, W. N., Febriana, B. W., Diniaty, A., & Fauzi'ah, L. (2018). Student profile in completing questions based on cognitive level of bloom's taxonomy by Anderson and Krathwohl. *AIP Conference Proceedings*, *2026*, 020063. https://doi.org/10.1063/1.5065023

Bahri, S., Simbolon, M., & Alhad, K. (2021). Development of Assessment Instrument of Scientific Literacy Ability for Students at Musamus

University. In *International Joined Conference on Social Science (ICSS 2021)* (pp. 395-398). Atlantis Press. https://doi.org/10.2991/assehr.k.211130.072
Bond, T. (2015). *Applying the Rasch Model*. Routledge. https://doi.org/10.4324/9781315814698
Cansız, N., & Cansız, M. (2019). Evaluating Turkish science curriculum with PISA scientific literacy framework. *Turkish Journal of Education*, *8*(3), 217–236. https://doi.org/10.19128/turje.545798
Darmana, A., Sutiani, A., Nasution, H. A., Ismanisa*, I., & Nurhaswinda, N. (2021). Analysis of Rasch Model for the Validation of Chemistry National Exam Instruments. *Jurnal Pendidikan Sains Indonesia*, *9*(3), 329–345. https://doi.org/10.24815/jpsi.v9i3.19618
Desnita, D., Festiyed, F., Novitra, F., Ardiva, A., & Yussavel Navis, M. (2022). The Effectiveness of CTL-based Physics E-module on the Improvement of the Creative and Critical Thinking Skills of Senior High School Students. *TEM Journal*, *11*(2), 802–810. https://doi.org/10.18421/TEM112-38
El Masri, Y. H., & Andrich, D. (2020). The Trade-Off between Model Fit, Invariance, and Validity: The Case of PISA Science Assessments. *Applied Measurement in Education*, *33*(2), 174–188. https://doi.org/10.1080/08957347.2020.1732384
Febrina, Bustami, U., & Asnawi, M. (2019). Analysis of Reading Comprehension Questions By Using Revised Bloom ' S Taxonomy on Higher Order. *English Education Journal*, *10*(1), 1–15. Retrieved from https://jurnal.usk.ac.id/EEJ/article/view/13253
Gani, U. A., Bambang, R., & Umam, K. (2018). Analisis Diskriminan Untuk Mengelompokkan Negara Maju Dan Negara Berkembang Dengan Metode Fishers Discriminant Analysis of Classify Developed Countries and Developing Countries With Fisher Method. *Jurnal Geuthèë: Penelitian Multidisiplin*, *01*(01), 1–12. https://doi.org/10.52626/jg.v1i1.6
Heuston, B. (2022). *Pursuing Excellence and Equity in Education: Vol. I* (PISA, pp. 637–664). OECD Publishing. https://doi.org/10.4018/978-1-7998-8649-5.ch026
Jamaluddin, J., Jufri, A. W., Ramdani, A., & Azizah, A. (2019). Profil Literasi Sains Dan Keterampilan Berpikir Kritis Pendidik Ipa Smp. *Jurnal Penelitian Pendidikan IPA*, *5*(1). https://doi.org/10.29303/jppipa.v5i1.185
Koerber, S., & Osterhaus, C. (2019). Individual Differences in Early Scientific Thinking: Assessment, Cognitive Influences, and Their Relevance for Science Learning. *Journal of Cognition and Development*, *20*(4), 510–533. https://doi.org/10.1080/15248372.2019.1620232
Kohar, D. (2022). Measuring the Effectiveness of the Brain-Based Learning Model on the Level of Reading Comprehension Based on Exposition Reading Structures in Junior High School. In *Educational Sciences: Theory and Practice* (Vol. 22, Issue 1, pp. 78–89). https://doi.org/10.12738/jestp.2022.1.0007
Lestari, H., Rahmawati, I., Siskandar, R., & Dafenta, H. (2021). Implementation of Blended Learning with A STEM Approach to Improve Student Scientific Literacy Skills During The Covid-19 Pandemic. *Jurnal Penelitian Pendidikan IPA*, *7*(2), 224. https://doi.org/10.29303/jppipa.v7i2.654
Lotulung Mareike, S. D., & Wullur Bill, G. (2017). Teacher Directive Discourse and Active Learning in Indonesian High Schools. *Journal of Modern Education Review*, *7*(3), 220–228. https://doi.org/10.15341/jmer(2155-7993)/03.07.2017/008
Marin, G., Sabogal, F., Marin, B. V., Otero-Sabogal, R., & Perez-Stable, E. J. (1987). Development of a Short Acculturation Scale for Hispanics. *Hispanic Journal of Behavioral Sciences*, *9*(2), 183–205. https://doi.org/10.1177/07399863870092005
Mujahidin, G. R., Rahman, G. G. A., Wilujeng, I., & Nugroho, S. D. (2023). Profil Literasi Sains Aspek Kompetensi Peserta Didik Melalui Pembelajaran Berbasis Budaya Lokal Reyog Ponorogo. *Jurnal Penelitian Pendidikan IPA*, *9*(2), 965–970. https://doi.org/10.29303/jppipa.v9i2.2753
Mulyanti, S., Suwahono, S., Setiowati, H., & Ningrum, L. S. (2022). Validity Analysis Using the Rasch Model in the Development of Alkane Concept Test Instruments. *Jurnal Penelitian Pendidikan IPA*, *8*(3), 1142–1147. https://doi.org/10.29303/jppipa.v8i3.1383
Muntazhimah, M., Putri, S., & Khusna, H. (2020). Rasch Model untuk Memvalidasi Instrumen Resiliensi Matematis Mahasiswa Calon Guru Matematika. *JKPM (Jurnal Kajian Pendidikan Matematika)*, *6*(1), 65. https://doi.org/10.30998/jkpm.v6i1.8144
Oktarina, R., Fitria, Y., Ahmad, S., & Zen, Z. (2023). *Development of STEM-Oriented E-Modules to Improve Science Literacy Ability of Elementary School Students*. *9*(1), 5460–5465. https://doi.org/10.29303/jppipa.v9i7.4503
Pratiwi, I. (2019). Efek Program Pisa Terhadap Kurikulum Di Indonesia. *Jurnal Pendidikan Dan Kebudayaan*, *4*(1), 51–71. https://doi.org/10.24832/jpnk.v4i1.1157
Rakkapao, S., Prasitpong, S., & Arayathanitkul, K. (2016). Analysis test of understanding of vectors

with the three-parameter logistic model of item response theory and item response curves technique. *Physical Review Physics Education Research*, *12*(2), 1–10. https://doi.org/10.1103/PhysRevPhysEducRes.12.020135

Rohmah, F., Rusilowati, A., Supriyadi, S., & Widowati, T. (2023). Validitas dan reliabilitas instrumen penilaian berbasis android untuk menilai kemampuan siswa SMK dalam mendesain busana secara digital. *JPPI (Jurnal Penelitian Pendidikan Indonesia)*, *9*(2). https://doi.org/10.29210/020232131

Rusilowati, A. (2018). Asesmen Literasi Sains: Analisis Karakteristik Instrumen dan Kemampuan Siswa Menggunakan Teori Tes Modern Rasch Model. *Prosiding Seminar Nasional Fisika Universitas Riau Ke-3*, *September*, 2–15. Retrieved from https://snf.fmipa.unri.ac.id/wp-content/uploads/2019/03/0.-300B-2-15NI.pdf

Sadhu, S., & Laksono, E. W. (2018). Development and validation of an integrated assessment for measuring critical thinking and chemical literacy in chemical equilibrium. *International Journal of Instruction*, *11*(3), 557–572. https://doi.org/10.12973/iji.2018.11338a

Sahnaz, S., & Kuswandi, P. C. (2023). Karangsari's Starfruit Agrotourism; Blitar's Local Pride as A Source of Biology Worksheet Apps to Enhance Students' Science Literacy. *Jurnal Penelitian Pendidikan IPA*, *9*(8), 5827-5833. https://doi.org/10.29303/jppipa.v9i8.3888

Setyorini, A. D., Yamtinah, S., Mahardiani, L., & Saputro, S. (2021). A Rasch analysis of item quality of the chemical literacy assessment for investigating student's chemical literacy on chemical rate concepts. *European Journal of Educational Research*, *10*(4), 1769–1779. https://doi.org/10.12973/EU-JER.10.4.1769

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*, *16*(S1), 1–5. https://doi.org/10.1007/s10763-018-9911-1

Sukmawati, W., & Zulherman, Z. (2023). Analysis of Changes in Students' Scientific Literacy Ability After Attending Lectures Using the RADEC Model. *Jurnal Penelitian Pendidikan IPA*, *9*(3), 1039–1044. https://doi.org/10.29303/jppipa.v9i3.2846

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Permodelan RASCH Pada Assesment Pendidikan* (I). Trim Komunikata.

Suparya, I. K., I Wayan Suastra, & Putu Arnyana, I. B. (2022). Rendahnya Literasi Sains: Faktor Penyebab dan Alternatif Solusinya. *Jurnal Ilmiah Pendidikan Citra Bakti*, *9*(1), 153–166. https://doi.org/10.38048/jipcb.v9i1.580

Susac, A., Planinic, M., Klemencic, D., & Milin Sipus, Z. (2018). Using the Rasch model to analyze the test of understanding of vectors. *Physical Review Physics Education Research*, *14*(2), 23101. https://doi.org/10.1103/PhysRevPhysEducRes.14.023101

Susongko, P., Arfiani, Y., & Kusuma, M. (2021). Determination of gender differential item functioning in tegal-students' scientific literacy skills with integrated science (Slisis) test using rasch model. *Jurnal Pendidikan IPA Indonesia*, *10*(2), 270–281. https://doi.org/10.15294/jpii.v10i2.26775

Tesio, L., Caronni, A., Kumbhare, D., & Scarano, S. (2023). Interpreting results from Rasch analysis 1. The "most likely" measures coming from the model. *Disability and Rehabilitation*, 1–13. https://doi.org/10.1080/09638288.2023.2169771

Wati, M., Husna, A., Miriam, S., & Mahtari, S. (2018). *Developing Cognitive Instruments Based on Science Literacy on Sound Materials*, *274*, 181–184. https://doi.org/10.2991/iccite-18.2018.41

Woudstra, A. J., Meppelink, C. S., Pander Maat, H., Oosterhaven, J., Fransen, M. P., & Dima, A. L. (2019). Validation of the short assessment of health literacy (SAHL-D) and short-form development: Rasch analysis. *BMC Medical Research Methodology*, *19*(1), 1–10. https://doi.org/10.1186/s12874-019-0762-4

Zahara, S. R., Imanda, R., & Alvina, S. (2022). Development of Assessment Sheet for Measuring Students' Scientific Literacy Level in the Era of Revolution 4.0. *Jurnal Penelitian Pendidikan IPA*, *8*(3), 1096–1101. https://doi.org/10.29303/jppipa.v8i3.1169