JPII 13 (2) (2024) 209-218

Jurnal Pendidikan IPA Indonesia

https://journal.unnes.ac.id/journals/jpii

ANALYSIS OF ITEMS FOR MEASURING TECHNOLOGICAL LITERACY USING THE RASCH MODEL

S. Jumini*[1], S. Madnasri*[2], E. Cahyono[3], Y. M. Nadrah[4], H. Hamzah[5], F. Fatiatun[6]

[1,6]Universitas Sains Al-Qur'an of Central Java in Wonosobo, Indonesia
[2]Departemen of Physiscs, Universitas Negeri Semarang, Indonesia
[3]Departemen of Chemistry, Universitas Negeri Semarang, Indonesia
[4]Departemen of Chemistry, Universitas Sains Islam Malaysia, Malaysia
[5]Madrasah Aliyah Negeri 2 Wonosobo, Indonesia

DOI: 10.15294/04k5q459

Accepted: May 16th, 2024. Approved: June 29th, 2024. Published: June 30th 2024

ABSTRACT

This study aims to describe the quality of the items on the instrument for measuring technological literacy in terms of validity, reliability, level of difficulty, and discrimination. This research used a descriptive quantitative approach with a construction and validation design. Data was collected through a test with 15 multiple-choice questions and a questionnaire with 40 questions. The test was conducted on 30 6th-semester physics students in the integrated science course. Data were analyzed by describing data processing results using the Rasch Model with Winsteps software. This research shows that the instrument to measure technological literacy has seven valid and six invalid questions regarding knowledge. Regarding attitude, 36 questions are valid, and four questions are invalid. The item reliability value for the knowledge is 0.93 in the very good category. For attitude, it is 0.98 in the excellent category. It shows that the questionnaire items are reliable. In terms of the level of difficulty, there are two easy, five medium, five difficult, and three highly difficult items for knowledge questions. Only students with high abilities can answer in the highly difficult question category. There are seven easy, 14 medium, nine difficult, and ten highly difficult items for attitude questions. It is concluded that these technological literacy test items are suitable for assessment and can contribute to measuring technological literacy, which has not been done much.

Keywords: analysis of test items; technological literacy; rasch model

INTRODUCTION

Excellent education cannot be separated from implementing teaching and learning activities with a good evaluation (Ferrando & Lorenzo-Seva, 2018; Alfarisa & Purnama, 2019). Evaluation is a data collection process that determines to what extent, in what terms, and how educational goals have been achieved (Arikunto, 2017; Keinänen et al., 2018). Learning evaluation is critical because it describes success in realizing educational goals. Evaluation is also valuable for decision-making. Evaluation includes measurement and assessment. Measurement is comparing something with one measure, while assessment is deciding whether something is good or bad using a measure. Assessment interprets measurement data based on specific criteria or rules (Widoyoko, 2012). Assessment is a qualitative activity, while measuring is a quantitative activity.

Assessment is essential to learning (Rahmawati et al., 2023). Assessment is beneficial for students in seeing how far they have succeeded in following the learning process and can motivate them (Nahijah et al., 2018; Alfarisa & Purnama, 2019; Azizah & Wahyuningsih, 2020). Meanwhile, for teachers, assessment can provide an overview of the progress of students' success, the accuracy of the material taught, and the methods

***Correspondence Address**
E-mail: srijumini@unsiq.ac.id;
sutiknomadnasri@mail.unnes.ac.id

used in learning. School evaluation results can be the basis for decision-making and policy-making and a benchmark for an educational unit's quality (Purniasari et al., 2021). Students' abilities are usually described by their total scores. If students reach the minimum passing grade or exceed the average, they are considered to have high ability, and vice versa. However, this raw score cannot measure students' achievements (Ayub et al., 2020; Purniasari et al., 2021). Therefore, decisions based on students' scores are not appropriate.

Teachers are the most responsible for evaluating learning in schools or education units. The role of teachers or educators is not only to plan and implement learning but also to assess (Susanta & Susanto, 2021). The assessment results show the quality of the learning. For this reason, teachers must have good skills in carrying out learning evaluations, planning, preparing assessment instruments, analysis, and interpretation. Teachers must also have skills in compiling good question items and analyzing them. A good assessment instrument must meet reliability criteria. Reliability criteria include validity, reliability, level of difficulty, and discrimination. An instrument trial is carried out to test the reliability or quality of the instrument. Test development and validation are crucial in assessment, the first and foremost process in evaluating educational management (Ismail & Zubairi, 2021; Ferrando et al., 2022).

Instrument trials are carried out internally and externally. An internal trial was conducted with experts to determine the validity of content, construct, and grammar (Almanasreh et al., 2019; Widoyoko, 2012). Expert judgment is needed to consider whether the instrument's structure and the scientific structure used in compiling the instrument are correct. To strengthen the validity of the assessment instrument, an external trial, called a field trial, must be carried out (Arikunto, 2012). Field trials are carried out on subjects that are equivalent to the subjects that will be assessed. After the trial, it is necessary to analyze the results, including validity, reliability, level of difficulty, and discrimination of the questions. Validity means the extent to which an assessment instrument carries out its assessment function accurately and reliably and the extent to which the test measures what it should measure (Sudaryono et al., 2013). Reliability is the extent to which the assessment results can be trusted. Several measurements on the same subject obtain relatively similar assessment results. The difficulty level of the assessment items shows the extent to which the subject accurately answers the questions. A good question is neither too difficult nor too easy. Meanwhile, discrimination shows the extent of students' abilities.

It is necessary to analyze the items for assessing learning outcomes to determine the validity, reliability, level of difficulty, and discrimination of the questions (Balsalobre-Fernández et al., 2016; Fuller et al., 2020). Items for assessing learning outcomes from field trials can be analyzed using the classical method, Classical Test Theory (CTT), or the modern method, Item Response Theory (IRT). Classical test theory is based on observed scores, the sum of the actual and measurement error scores. The quality of the test items is determined by the level of difficulty and discrimination (Hardianti et al., 2021). However, the characteristics of inconsistent test items depend on the students' abilities (Erfan et al., 2020). Students' abilities will appear low if the test questions are difficult, whereas if the test questions are easy, students' abilities will appear high. Weaknesses of classical test theory include: 1) the level of difficulty and discrimination of the questions depends on the test takers; 2) it analyzes by comparing the abilities of test takers in the upper, middle, and lower groups; 3) the concept of score reliability is defined in terms of parallel tests; 4) Determining how students obtain tests according to their abilities does not have a supporting theory; 5) Standard Error of Measurement (SEM) applies to all test takers (Alfarisa & Purnama, 2019).

The modern method is Item Response Theory (IRT), such as the Rasch Model (one-parameter logistic model). The Rasch Model was first discovered by Dr. Georg Rasch, a mathematician from Denmark, who was here to overcome the weaknesses of the classical method. The Rasch Model uses raw scores in different ways to produce measurement scales with equal intervals to provide accurate information about test takers and the quality of the questions (Sumintono & Widhiarso, 2014). It can provide information related to test takers and the quality of the questions they work on because it forms a measurement ratio with the same distance. It does not have discrimination and guess parameters. The level of difficulty of the items affects the performance of a test, so the selection of items in preparing the test affects the quality of the test according to the needs and objectives of the assessment. The Rasch Model can predict missing data and produce appropriate statistical analysis in research.

This study aims to describe the quality of the items on the instrument for measuring technological literacy in the aspects of validity, reliability, level of difficulty, and discrimination, as well as the distribution of students' technological literacy through modern methods using Item Response Theory (IRT), which is the Rasch model. This study analyzes test instruments measuring technological literacy. Technological literacy is understanding, assessing, managing, and using technology (aher & Abtaria, 2017; Asrizal et al., 2018). Technological literacy includes knowledge, ability, critical thinking, and decision-making (Zainurrisalah et al., 2018). The following are taxonomy of technological literacy: 1) Medical Technology, technology related to diagnosing, treating, and preventing disease and other damage to the body or mind (Yan et al., 2018); 2) Agriculture and Related Biotechnology, technology related to the cultivation of plants and animals for food, feed, fiber, fuel, or other purposes (Pfeifer et al., 2021), 3) Energy and Power Technology, technology related to the utilization of energy resources and converting energy into power (Ahmadi et al., 2018); 4) Information and Communication Technology, technology which includes educational technology, developed to collect, manipulate, classify, store and retrieve information (Upadhayaya, 2023); 5) Transportation Technology, technological processes and systems in which people or goods are moved from one place to another (Wensveen, 2023); 6) Manufacturing Technology, technological processes and systems that convert materials into finished products (Diegel et al., 2019); 7) Construction Technology, technological processes, and systems related to the construction of buildings, roads, embankments, and other structures (Asunda, 2012). Someone with technological literacy will understand the nature of technology, criticize it, and use it well to achieve goals in life (Raharjo et al., 2014). Good quality items are needed as a measurement instrument to measure technological literacy accurately.

METHODS

This research used a descriptive quantitative approach with a construction and validation design (Creswell & Poth, 2018) to analyze the difficulty level, validity, and reliability of the test items and the level of students' abilities. The instrument for measuring technological literacy is based on knowledge, ability, critical thinking, and decision-making. Knowledge includes (1) recognizing the pervasiveness of technology in everyday life, (2) understanding basic engineering concepts and terms, such as systems, constraints, and trade-offs, (3) being familiar with the nature and limitations of the engineering design process, (4) knowing how technology has shaped human history and how humans have; (5) recognizing that all technology carries risks and only some can be anticipated, (6) appreciating that the development and use of technology involve trade-offs and balancing costs and benefits; (6) understanding that technology reflects the values and culture of society.

Critical thinking and decision-making include (1) asking questions about the benefits and risks of technology, (2) systematically weighing available information about the benefits, risks, costs, and trade-offs of technology, and (3) participating, as necessary, in decisions about technology development and use. Ability includes: (1) having a variety of hands-on skills, such as operating various home and office equipment and using a computer for word processing and surfing the internet, (2) identifying and repairing simple mechanical or technological problems at home or work, (3) applying basic mathematical concepts related to probability, scale, and estimation to make informed judgments about the risks and benefits of technology, (4) using the design-thinking process to solve problems faced in everyday life, (5) obtaining information about technological issues of concern from various sources (Garmire & Pearson, 2006).

Technological literacy is measured in five dimensions (Groth et al., 2007; Hohlfeld et al., 2010): 1) Technology Knowledge, including technological developments, the meaning and content of technology, basic principles of technology, essence, and scope of technology (Willis et al., 2019); 2) Technology Application, including daily applications, technological knowledge applications, problem-solving, and evaluation; 3) Technology Attitude, including an interest in technology products, honesty, initiative, objectivity, and the like; 4) Technology Estimation, including estimating technology, selecting appropriate technology, and assessing and evaluating technology products; 5) Technology Resources, including human resources, time, raw materials, energy, information, and costs (Yu-Te et al., 2014; Neumeyer et al., 2021).

The measurement of students' technological literacy in this research includes knowledge, thinking, and attitude. Ability is not measured due to many limitations. Knowledge is measured with 15 multiple-choice questions, while attitude is measured by examining the technology literacy profile through a 40-item questionnaire. After experts validated it, the instrument was implemented to measure the technological literacy of

UNSIQ Physics Education students. The sampling technique is saturated, taking one class as the research subject. The research subjects in the field trial were 30 6th-semester physics students in integrated science courses. Data were analyzed by describing data processing results using the Rasch Model with Winsteps software.

RESULTS AND DISCUSSION

This research describes the quality of the questions in terms of validity, reliability, and level of difficulty according to the Rasch model. The validity measured in this research is content validity, which examines the questions' material, construction, and language quality (Schaufeli et al., 2020; Jumini et al., 2023). Content validity was measured using the CVR method, involving seven material experts, two science learning experts, curriculum experts, learning evaluation experts, and two industry practitioners. The index proposed by Putranta et al. (2019) was used to determine the level of expert agreement on internal validity. The results were analyzed with AikenV. If an item has a V index ≤ 0.4, the validity is low (LV); if 0.4 ≤ V ≤ 0.8, the validity is medium (MV); if ≥ 0.8, the validity is high (HV). The results of internal validation from the seven validators of the technology literacy measurement instrument are presented in Table 1.

Table 1. Results of Internal Validity Assessment of Technology Literacy Question Items

No	ASPECT	V	CRITERIA
1.	Completeness of the question framework	0.9	HV
2.	Legibility	0.8	HV
3.	Following the rules	0.9	HV
4.	Effective and efficient language	0.9	HV
5.	Graphic components	0.7	MV
6.	Assessment keys and scores	0.9	HV
7.	Display of the tested ability	0.9	HV

After the expert declared the question items valid, an external test was carried out, and the results were analyzed using the Rasch Model.

1. Validity of Technological Literacy Question Items. Validity shows the extent to which a measuring instrument is accurate and precise in its measurement function (Azwar, 2019; Peeters & Harpe, 2022). Careful means detecting minor differences in the measured attributes. Test questions are highly valid if the questions provide appropriate results according to the test objectives. A questionnaire is valid if it reveals what it measures. The quality of validity of the items in the Rasch Model is achieved if they fulfill several criteria (Sumintono & Widhiarso, 2015; Peeters & Augustine, 2023).

a. The Outfit MNSQ (Mean Square) value received is 0.5 < Outfit – MNSQ < 1.5; b. The Outfit ZSTD (Z – Standard) value accepted is: –2.0 < ZSTD < +2.0; c. Pt Measure Corr (Point Measure Correlation) value: 0.4 < Point Measure Corr < 0.85.

Based on the MNSQ and ZSTD outfit value criteria and Point Measure Correlation, valid questions are number 1, 4, 6, 7, 9, 10, and 13. In comparison, invalid questions are numbers 2, 3, 5, 8, 11, 12, 14 and 15. The results of the analysis of the 15 questions are presented in Figure 1.

TABLE 13.1 Analisis Soal Tes Literasi Teknologi ZOU171WS.TXT May 1 2023 1:41
INPUT: 33 Person 15 Item REPORTED: 33 Person 15 Item 2 CATS MINISTEP 4.3.2

Person: REAL SEP.: 1.23 REL.: .60 ... Item: REAL SEP.: 3.22 REL.: .91

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
7	0	33	6.02	1.85	MAXIMUM MEASURE				.00	.00	100.0	100.0	S7
1	5	33	2.88	.61	.90	-.10	1.46	.78	.54	.56	93.8	89.2	S1
13	5	33	2.88	.61	.62	-.91	.31	-1.00	.75	.56	93.8	89.2	S13
10	6	33	2.54	.55	1.11	.41	.81	-.10	.53	.55	84.4	87.0	S10
4	8	33	2.01	.49	.65	-1.27	.56	-.94	.71	.53	87.5	82.5	S4
6	8	33	2.01	.49	1.15	.59	.98	.11	.48	.53	81.3	82.5	S6
9	12	33	1.21	.42	1.10	.58	1.02	.16	.45	.50	68.8	74.0	S9
5	19	33	.09	.39	1.12	.88	2.42	3.37	.29	.43	65.6	68.0	S5
2	25	33	-.92	.44	.97	-.08	.81	-.20	.38	.35	75.0	77.1	S2
15	25	33	-.92	.44	1.05	.28	.96	.10	.32	.35	81.3	77.1	S15
14	27	33	-1.35	.49	.75	-.90	.49	-.71	.48	.32	84.4	82.5	S14
8	29	33	-1.90	.57	.89	-.18	.66	-.17	.34	.27	90.6	88.0	S8
3	30	33	-2.26	.64	1.05	.25	.66	-.10	.26	.25	87.5	90.6	S3
11	31	33	-2.74	.76	1.24	.57	3.44	1.74	-.05	.21	93.8	93.6	S11
12	32	33	-3.52	1.04	1.13	.44	1.93	.99	.00	.16	96.9	96.8	S12
MEAN	17.5	33.0	.40	.65	.98	.0	1.18	.3			84.6	84.2	
P.SD	11.1	.0	2.56	.36	.19	.6	.84	1.1			9.2	7.7	

TABLE 13.3 Analisis Soal Tes Literasi Teknologi ZOU171WS.TXT May 1 2023 1:41
INPUT: 33 Person 15 Item REPORTED: 33 Person 15 Item 2 CATS MINISTEP 4.3.2

Figure 1. Validity Analysis of Technological Literacy Test Questions

The analysis of test items to measure technological literacy in the attitude aspect is presented in Figure 2. Four questions are invalid based on the criteria for the MNSQ outfit and the ZSTD outfit (numbers 1, 6, 10, and 14). Meanwhile, according to the Point Measure Correlation criteria, all questions are valid.

```
        Item STATISTICS:  MEASURE ORDER

-------------------------------------------------------------------------------------------------
|ENTRY   TOTAL  TOTAL           MODEL|   INFIT    |  OUTFIT   |PTMEASUR-AL|EXACT  MATCH|       |
|NUMBER  SCORE  COUNT  MEASURE  S.E. |MNSQ  ZSTD|MNSQ  ZSTD|CORR.  EXP.| OBS%  EXP%| Item  |
|------------------------------------+----------+----------+-----------+-----------+-------|
|    26     42     31     3.65    .36|1.06   .33|1.08   .40| -.25   .19| 54.8  66.0| A26   |
|    10     46     31     3.18    .33|1.47  1.75|1.52  1.97|  .12   .20| 61.3  56.7| A10   |
|    18     49     31     2.87    .31| .99   .05|1.01   .11| -.02   .21| 48.4  53.4| A18   |
|    36     49     31     2.87    .31| .75  -.96| .76  -.95|  .11   .21| 54.8  53.4| A36   |
|    24     50     31     2.77    .31|1.16   .63|1.17   .69| -.06   .22| 51.6  53.9| A24   |
|    20     54     31     2.41    .29| .98   .04|1.00   .08|  .16   .23| 54.8  57.6| A20   |
|    40     54     31     2.41    .29| .94  -.09| .84  -.47|  .20   .23| 64.5  57.6| A40   |
|    28     56     31     2.25    .28| .63 -1.27| .63 -1.32| -.07   .24| 74.2  59.7| A28   |
|    32     58     31     2.09    .27|1.06   .30|1.03   .19|  .04   .25| 64.5  61.2| A32   |
|    38     58     31     2.09    .27| .99   .07| .98   .03| -.13   .25| 71.0  61.2| A38   |
|     2     61     31     1.88    .26| .86  -.36| .80  -.58|  .38   .26| 71.0  61.4| A2    |
|    22     68     31     1.45    .23| .92  -.20| .85  -.42|  .15   .29| 64.5  58.6| A22   |
|    30     68     31     1.45    .23|1.58  1.82|1.48  1.52|  .01   .29| 51.6  58.6| A30   |
|    12     69     31     1.39    .23|1.12   .51|1.04   .23|  .34   .29| 48.4  58.2| A12   |
|    14     82     31      .79    .20|1.93  3.49|1.90  3.30|  .34   .32| 19.4  37.1| A14   |
|    34     84     31      .71    .20|1.23  1.09|1.21   .98|  .39   .32| 41.9  33.1| A34   |
|     8     93     31      .37    .19| .92  -.35| .90  -.45|  .43   .31| 25.8  28.2| A8    |
|     4     96     31      .26    .19| .96  -.16| .99   .01|  .24   .31| 45.2  28.4| A4    |
|    29     98     31      .18    .19|1.23  1.18|1.31  1.52|  .20   .31| 12.9  28.4| A29   |
|     6    110     31     -.29    .21| .51 -2.55| .54 -2.25|  .63   .28| 58.1  41.2| A6    |
|    23    110     31     -.29    .21| .59 -2.02| .68 -1.44|  .16   .28| 38.7  41.2| A23   |
|    37    120     31     -.76    .23| .58 -1.66| .59 -1.61|  .27   .25| 67.7  57.9| A37   |
|    21    123     31     -.93    .24|1.13   .52|1.17   .65|  .31   .24| 41.9  59.6| A21   |
|    16    125     31    -1.05    .25| .84  -.44| .86  -.39|  .04   .23| 77.4  59.6| A16   |
|    11    126     31    -1.12    .25| .76  -.75| .72  -.92|  .39   .23| 64.5  59.5| A11   |
|    15    126     31    -1.12    .25| .62 -1.32| .58 -1.54|  .16   .23| 77.4  59.5| A15   |
|     3    129     31    -1.32    .27| .67 -1.11| .68 -1.07|  .61   .22| 61.3  58.8| A3    |
|    35    131     31    -1.47    .28| .64 -1.25| .65 -1.24|  .34   .22| 61.3  57.5| A35   |
|     7    132     31    -1.55    .28| .62 -1.34| .65 -1.25|  .18   .21| 74.2  57.0| A7    |
|    39    132     31    -1.55    .28|1.10   .43|1.09   .41|  .43   .21| 58.1  57.0| A39   |
|    33    134     31    -1.71    .29|1.44  1.39|1.47  1.52|  .37   .21| 64.5  55.5| A33   |
|     1    136     31    -1.89    .30|1.56  1.77|1.58  1.86|  .41   .20| 51.6  53.2| A1    |
|    25    137     31    -1.98    .31| .84  -.52| .80  -.68|  .65   .20| 71.0  53.9| A25   |
|     5    138     31    -2.07    .31|1.00   .10| .99   .07|  .39   .19| 58.1  53.5| A5    |
|    27    138     31    -2.07    .31| .91  -.25| .92  -.23|  .15   .19| 58.1  53.5| A27   |
|    31    140     31    -2.28    .32| .77  -.88| .76  -.93|  .25   .19| 54.8  55.0| A31   |
|    13    143     31    -2.61    .35| .87  -.46| .87  -.43|  .15   .17| 58.1  62.1| A13   |
|    17    144     31    -2.73    .36| .84  -.57| .82  -.66|  .29   .17| 61.3  65.6| A17   |
|    19    144     31    -2.73    .36| .79  -.79| .77  -.88|  .42   .17| 67.7  65.6| A19   |
|     9    149     31    -3.52    .45| .96   .00| .97   .03|  .09   .13| 80.6  80.9| A9    |
|------------------------------------+----------+----------+-----------+-----------+-------|
| MEAN   100.0   31.0      .00    .28| .97   -.1| .97   -.1|           | 57.2  54.5|       |
| P.SD    36.1     .0     1.98    .05| .30   1.1| .30   1.1|           | 14.7  10.8|       |
-------------------------------------------------------------------------------------------------
```

Figure 2. Validity Analysis of Technology Literacy Profile Questionnaire Questions

2. Reliability of Technological Literacy Test Items. Reliability is the consistency of a series of measuring instruments (Borges & Driller, 2016; Arikunto, 2017). Reliability always shows the condition of research instruments in various forms, such as the same test results if carried out by different people (inter-rater), the same test results if carried out by the same person at different times (retesting), the same test results if carried out by different people at the same time with different tests (parallel form), and the same test results using various constructive questions (internal consistency). The criteria for determining Item and Person Reliability are presented in Table 2 (Widhiarso, 2014; Silvia et al., 2021).

Table 2. Reliability Criteria in Rasch Modeling

Reliability Value (Person/Item)	Interpretation
>0,94	Excellent
0,91 – 0,94	Very good
0,81 – 0,90	Good
0,67 – 0,80	Average
< 0,67	Poor

The results of the reliability test for 15 questions are shown in Figure 3. The person reliability value is 0.67 in the average category. It shows students' consistency when answering questions in fairly consistent conditions. The item reliability value is 0.93 in the very good category, indicating that the quality of the questions in the instrument has an outstanding reliability aspect. From the Alpha Cronbach (KR-20) formula, 0.63 in the medium category was obtained, indicating that the overall reliability of the instrument is moderate.

```
     SUMMARY OF 33 MEASURED (EXTREME AND NON-EXTREME) Person
-------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT         OUTFIT    |
|          SCORE      COUNT    MEASURE   S.E.       MNSQ   ZSTD    MNSQ   ZSTD |
|------------------------------------------------------------------------------|
| MEAN       7.9       15.0        .60     .82                                 |
|  SEM        .4         .0        .26     .03                                 |
| P.SD       2.2         .0       1.46     .19                                 |
| S.SD       2.2         .0       1.48     .19                                 |
| MAX.      14.0       15.0       5.35    1.85                                 |
| MIN.       3.0       15.0      -2.35     .76                                 |
|------------------------------------------------------------------------------|
| REAL RMSE     .92 TRUE SD    1.14  SEPARATION  1.23  Person RELIABILITY  .60 |
|MODEL RMSE     .84 TRUE SD    1.19  SEPARATION  1.41  Person RELIABILITY  .67 |
| S.E. OF Person MEAN = .26                                                    |
-------------------------------------------------------------------------------
Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .63  SEM = 1.35

     SUMMARY OF 14 MEASURED (NON-EXTREME) Item
-------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT         OUTFIT    |
|          SCORE      COUNT    MEASURE   S.E.       MNSQ   ZSTD    MNSQ   ZSTD |
|------------------------------------------------------------------------------|
| MEAN      18.7       33.0        .00     .57        .98    .04   1.18    .29 |
|  SEM       2.9         .0        .60     .04        .05    .17    .23    .31 |
| P.SD      10.4         .0       2.15     .16        .19    .63    .84   1.11 |
| S.SD      10.8         .0       2.23     .17        .19    .65    .87   1.16 |
| MAX.      32.0       33.0       2.88    1.04       1.24    .88   3.44   3.37 |
| MIN.       5.0       33.0      -3.52     .39        .62  -1.27    .31  -1.00 |
|------------------------------------------------------------------------------|
| REAL RMSE     .61 TRUE SD    2.06  SEPARATION  3.36  Item   RELIABILITY  .92 |
|MODEL RMSE     .59 TRUE SD    2.07  SEPARATION  3.51  Item   RELIABILITY  .93 |
| S.E. OF Item MEAN = .60                                                      |
-------------------------------------------------------------------------------
```

Figure 3. Reliability Analysis of Technological Literacy Test Questions

The reliability analysis of questionnaire items to measure technological literacy in the attitude aspect is presented in Figure 4. The person reliability value is 0.59 in the weak category. When answering questionnaire questions, students are inconsistent. The item reliability value was 0.98 in the excellent category, indicating that the questionnaire items are reliable. The reliability value using the Alpha Cronbach (KR-20) formula was 0.58 in the medium category, indicating moderate consistency in the overall question structure.

```
|-----------------------------------------------------------------------------|
| MEAN     129.1      40.0          .41     .23        .99   -.15    .97   -.19 |
|  SEM       1.2        .0          .07     .00        .07    .27    .06    .23 |
| P.SD       6.8        .0          .36     .00        .37   1.46    .31   1.27 |
| S.SD       6.9        .0          .37     .00        .38   1.49    .31   1.29 |
| MAX.     149.0      40.0         1.50     .24       2.09   3.48   1.66   2.35 |
| MIN.     118.0      40.0         -.17     .23        .44  -2.86    .48  -2.60 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .25 TRUE SD     .26  SEPARATION  1.07  Person RELIABILITY  .54 |
|MODEL RMSE    .23 TRUE SD     .28  SEPARATION  1.21  Person RELIABILITY  .59 |
| S.E. OF Person MEAN = .07                                                   |
-------------------------------------------------------------------------------
Person RAW SCORE-TO-MEASURE CORRELATION = 1.00
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .58  SEM = 4.39

     SUMMARY OF 40 MEASURED Item
-------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT         OUTFIT     |
|          SCORE      COUNT   MEASURE    S.E.       MNSQ   ZSTD   MNSQ   ZSTD  |
|-----------------------------------------------------------------------------|
| MEAN     100.0      31.0          .00     .28        .97   -.09    .97   -.10 |
|  SEM       5.8        .0          .32     .01        .05    .18    .05    .18 |
| P.SD      36.1        .0         1.98     .05        .30   1.15    .30   1.13 |
| S.SD      36.6        .0         2.00     .06        .31   1.16    .31   1.14 |
| MAX.     149.0      31.0         3.65     .45       1.93   3.49   1.90   3.30 |
| MIN.      42.0      31.0        -3.52     .19        .51  -2.55    .54  -2.25 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .29 TRUE SD    1.95  SEPARATION  6.64  Item   RELIABILITY  .98 |
|MODEL RMSE    .28 TRUE SD    1.96  SEPARATION  6.92  Item   RELIABILITY  .98 |
| S.E. OF Item MEAN = .32                                                     |
-------------------------------------------------------------------------------
Item RAW SCORE-TO-MEASURE CORRELATION = -.99
Global statistics: please see Table 44.
UMEAN=.0000 USCALE=1.0000
```

Figure 4. Results of Reliability Analysis of Technology Literacy Questionnaire Questions

3. Level of Difficulty of Technological Literacy Test Items. The difficulty level of a question item shows how many respondents are likely to answer a question item correctly (Arikunto, 2017; Andrade, 2018; Hamdu et al., 2020). The difficulty level of a question item states whether the question item is in the difficult, average, or easy category. A good question is neither too difficult nor too easy. The level of difficulty of the items in Rasch modeling is categorized based on Measure logit and the Standard Deviation (SD) value of the logit items and divided into four categories as shown in Table 3 (Planinic et al., 2019; Widhiarso, 2015). A high logit value indicates the highest level of difficulty of the questions.

Table 3. Criteria for Difficulty Levels of Question Items with Rasch Modeling

Measure Value (logit)	Interpretation of Question Item Difficulty
Measure logit < – SD logit	Easy items
– SD logit ≤ Measure logit ≤ 0	Medium items
0 ≤ Measure logit ≤ SD logit	Difficult items
Measure logit > SD logit	Highly Difficult Items

The analysis of the difficulty level of 15 technological literacy questions is summarized in Table 4.

Table 4. Level of Difficulty of Technological Literacy Test Items

Measure Value (logit)	Interpretation	Question Items
Measure logit < – 2,56	Easy items	11,12
– 2,56≤ Measure logit ≤ 0	Medium items	2,3,8,14,15
0 ≤ Measure logit ≤ 2,56	Difficult items	4,5,6,9,10
Measure logit > 2,56	Highly Difficult Items	1,7,13

The difficulty level of the test items measuring technological literacy can be confirmed in the item map shown in Figure 5. From this item map, questions 1, 7, and 13 are in the highest logit category, which is highly difficult to do, and only three students answered them correctly. Questions 4, 5, 6, 9, and 10 were difficult, but more students answered correctly. Questions 2, 3, 8, 14, and 15 are in the medium category, and 11 and 12 are in the easy category. All students answered them correctly.

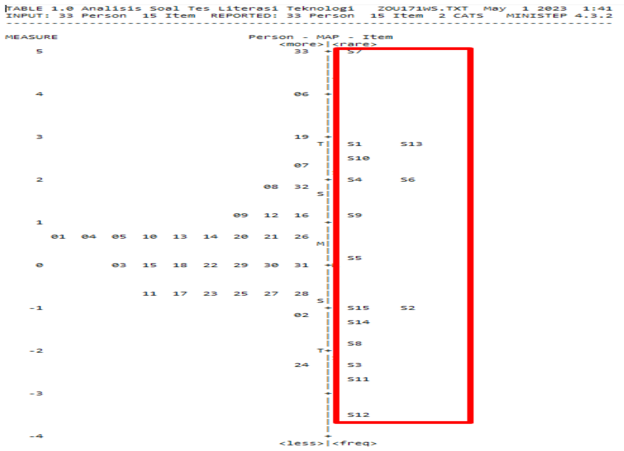

Figure 5. Item Map of Technological Literacy Analysis

Analysis of the difficulty level of the questionnaire items to measure technological literacy in the attitude aspect is presented in Table 5.

Table 5. Level of Difficulty of Technological Literacy Questionnaire Items

Measure Value (logit)	Interpretation	Amount	Question Items
Measure logit < – 1,98	Easy items	7	5,9,13,17,19,27,31
– 1,98≤ Measure logit ≤ 0	Medium items	14	1,3,6,7,11,15,16,21, 23,25,33,35,37,39
0 ≤ Measure logit ≤ 1,98	Difficult items	9	2,4,8,12,14,22,29,30,34
Measure logit > 1,98	Highly Difficult Items	10	10,18,20,24,26,28, 32,38,38,40

The difficulty level of the test items to measure technological literacy can be confirmed in the item map in Figure 6. From the item map, ten questions are in the highest logit category (10, 18, 20, 24, 26, 28, 32, 38, 38, and 40). They were very difficult to agree with, and no students agreed to the items. Next are nine questions (2, 4, 8, 12, 14, 22, 29, 30, and 34) in the difficult category. These questions were also difficult for students to agree with. Then, 14 questions were in the medium category, and seven were in the easy category; not all were approved by students. This result is in line with Hamdu et al. (2020).

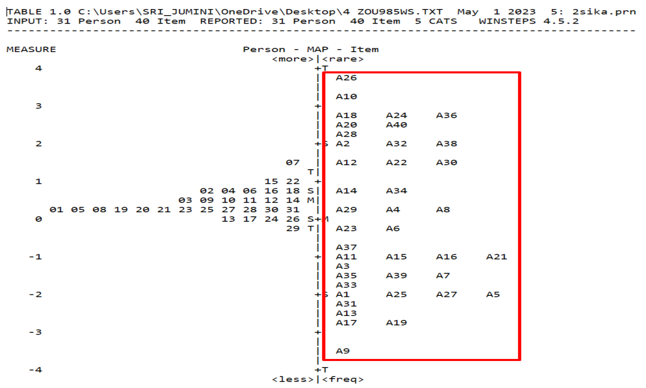

Figure 6. Questionnaire Item Map of Technological Literacy Analysis

In general, the quality of the technological literacy items is quite good. When used to measure technological literacy, the items accurately measure and express technological literacy in aspects of knowledge, critical thinking, and decision-making. They show the same technological literacy results if used to measure one competency at different times. These technological literacy questions also have a good difficulty level because not all questions are too difficult or too easy. It does not mean students have low abilities, so they cannot answer them, and vice versa. The criteria for the questions are complex, and even students with high abilities cannot answer them. Thus, the technological literacy test items are suitable for assessing technological literacy. An excellent technological literacy measuring tool provides an accurate picture of one's technological literacy. Someone with good technological literacy has excellent skills (Supriyadi et al., 2020; Triana et al., 2020).

The results of this research benefit the education system, especially in learning evaluation. There are few instruments for measuring technological literacy because it is not easy to measure, so it is rarely measured. This instrument provides an overview of how to define technology and technological literacy. It gives educational standards related to technological literacy and its curriculum, which is relatively new and has not been widely tested. This technological literacy measurement instrument also provides an overview of the interconnection between scientific and transdisciplinary disciplines. The results of this research can be used as a reference for how 21st-century learning is closely related to technology in various dimensions so that learning actors and education managers can develop similar instruments with different materials and sub-concepts.

CONCLUSION

This research shows that the instrument to measure technological literacy has seven valid and six invalid questions regarding knowledge. Regarding attitude, 36 questions are valid, and four questions are invalid. The item reliability value for the knowledge is 0.93 in the very good category. In comparison, for attitude, it is 0.98 in the excellent category. It shows that the questionnaire items are reliable. In terms of the level of difficulty, there are two easy, five medium, five difficult, and three highly difficult items for knowledge questions. Only students with high abilities can answer in the highly difficult question category. There are seven easy, 14 medium, nine difficult, and ten highly difficult items for attitude questions. Highly difficult items are in the highest logit category and difficult to agree on. No students agreed to the highly difficult and difficult question items. In comparison, students did not agree with all of the questions in the medium category and seven items in the easy category.

ACKNOWLEDGMENTS

Thanks are expressed to the Direktorat Jenderal Pendidikan Tinggi Riset dan Teknologi, Kementrian Pendidikan dan Kebudayaan Republik Indonesia for funding this research through the Doctoral Dissertation Research Grant Program with contract number Nomor Kontrak: 14.13.4/ UN37/PPK10/2023. Thanks are also conveyed to the Islamic University of Science Malaysia for being willing to help validate this instrument, so that this research can run well.

REFERENCES

Ahmadi, M. H., Ghazvini, M., Sadeghzadeh, M., Alhuyi Nazari, M., Kumar, R., Naeimi, A., & Ming, T. (2018). Solar power technology for electricity generation: A critical review. *Energy Science and Engineering*, *6*(5), 340–361.

Alfarisa, F., & Purnama, D. N. (2019). *Analisis Butir Soal Ulangan Akhir Semester Mata Pelajaran Ekonomi SMA Menggunakan RASCH Model*, *11*(2), 366-374.

Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy*, *15*(2), 214–221.

Andrade, C. (2018). Internal, External, and Ecological Validity in Research Design, Conduct, and Evaluation. *Indian Journal of Psychological Medicine*, *40*(5), 498–499.

Arikunto, S. (2012). *Dasar-dasar evaluasi Pendidikan* (Restu Damayanti (ed.); Edisi ke-2). Bumi Aksara.

Arikunto, S. (2017). *Pengembangan Instrumen Penelitian dan Penilaian Program*. Yokyakarta: Pustaka Pelajar.

Asrizal, A., Amran, A., Ananda, A., & Festiyed, F. (2018). Effectiveness of Adaptive Contextual Learning Model of Integrated Science by Integrating Digital Age Literacy on Grade VIII Students. *IOP Conference Series: Materials Science and Engineering*, *335*(1).

Asunda, P. A. (2012). Standards for technological literacy and STEM education delivery through career and technical education programs. *Journal of Technology Education*, *23*(2), 44–60.

Ayub, M. R. S. S. N., Istiyono, E., Munadi, S., Permadi, C., Pattiserlihun, A., & Sudjito, D. N. (2020). Analisa Penilaian Soal Fisika Menggunakan Model Rasch Dengan Program R. *Jurnal Sains Dan Edukasi Sains*, *3*(2), 46–52.

Azizah, A., & Wahyuningsih, S. (2020). Penggunaan Model Rasch Untuk Analisis Instrumen Tes Pada Mata Kuliah Matematika Aktuaria. *JUPITEK: Jurnal Pendidikan Matematika*, *3*(1), 45–50.

Azwar, S. (2019). *Penyusunan Skala Psikologi*. Yogyakarta: Pustaka Pelajar.

Balsalobre-Fernández, C., Kuzdub, M., Poveda-Ortiz, P., & Campo-Vecino, J. del. (2016). Validity and Reliability of the PUSH Wearable Device to Measure Movement Velocity During the Back Squat Exercise. *Journal of Strength and Conditioning Research*, *30*(7), 1968–1974.

Borges, N. R., & Driller, M. W. (2016). Wearable Lactate Threshold Predicting Device is Valid and Reliable in Runners. *Journal of Strength and Conditioning Research*, *30*(8), 2212–2218.

Creswell, J., & Poth, C. (2018). *Qualitative Inguiry Research Design: Choosing amoung five approaches* (4th ed.). California : SAGE Publications, Inc., Thousand Oaks.

Diegel, O., Nordin, A., & Motte, D. (2019). Additive Manufacturing Technologies. *Springer Series in Advanced Manufacturing*, 19–39.

Erfan, M., Maulyda, M. A., Hidayati, V. R., Astria, F. P., & Ratu, T. (2020). Tes Klasik Dan Model Rasch. *Indonesian Journal of Educational Research and Review*, *3*(1), 11–19.

Ferrando, P. J., & Lorenzo-Seva, U. (2018). Assessing the Quality and Appropriateness of Factor Solutions and Factor Score Estimates in Exploratory Item Factor Analysis. *Educational and Psychological Measurement*, *78*(5), 762–780.

Ferrando, P. J., Lorenzo-Seva, U., Hernández-Dorado, A., & Muñiz, J. (2022). Decalogue for the Factor Analysis of Test Items. *Psicothema*, *34*(1), 7–17.

Fuller, D., Colwell, E., Low, J., Orychock, K., Ann Tobin, M., Simango, B., Buote, R., van Heerden, D., Luan, H., Cullen, K., Slade, L., & Taylor, N. G. A. (2020). Reliability and Validity of Commercially Available Wearable Devices for Measuring Steps, Energy Expenditure, and Heart Rate: Systematic Review. *JMIR MHealth and UHealth*, *8*(9), 1–23.

Garmire, E., & Pearson, G. (2006). Tech tally: Approaches to assessing technological literacy. In *Tech Tally: Approaches to Assessing Technological Literacy*.

Groth, L. A., Dunlap, K. L., & Kidd, J. K. (2007). Becoming technologically literate through technology integration in PK-12 preservice literacy courses: Three case studies. *Reading Research and Instruction*, *46*(4), 363–386.

Hamdu, G., Fuadi, F. N., Yulianto, A., & Akhirani, Y. S. (2020). Items Quality Analysis Using Rasch Model To Measure Elementary School Students' Critical Thinking Skill On Stem Learning. *JPI (Jurnal Pendidikan Indonesia)*, *9*(1), 61.

Hardianti, H., Liliawati, W., & Tayubi, Y. R. (2021). Karakteristik Tes Kemampuan Berpikir Kritis Siswa SMA Pada Materi Momentum dan Impuls: Perbandingan Classical Theory Test (CTT) dan model Rasch. *WaPFi (Wahana Pendidikan Fisika)*, *8*(1), 21-28.

Hohlfeld, T. N., Ritzhaupt, A. D., & Barron, A. E. (2010). Development and Validation of the Student Tool for Technology Literacy (ST 2 L). *Journal of Research on Technology in Education*, *42*(4), 361–389.

Ismail, F. K. M., & Zubairi, A. M. B. (2021). Item Objective Congruence Analysis for Multidimensional Items Content Validation of a Reading Test in Sri Lankan University. *English Language Teaching*, *15*(1), 106.

Jumini, S., Madnasri, S., Cahyono, E., & Parmin, P. (2023). Analisis Kualitas Butir Soal Pengukuran Literasi Sains Melalui Teori Tes Klasik Dan Rasch Model. *Prosiding Seminar Nasional Pascasarjana UNNES*, 758–765.

Keinänen, M., Ursin, J., & Nissinen, K. (2018). How to measure students' innovation competences in higher education: Evaluation of an assessment tool in authentic learning environments. *Studies in Educational Evaluation*, *58*, 30–36.

Nahijah, A. R., Serevina, V., & Delina, M. (2018). The Development of High Order Thinking Skills (HOTS) Assessment Instrument for Temperature and Heat Learning. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, *4*(1), 19-26.

Neumeyer, X., Santos, S. C., & Morris, M. H. (2021). Overcoming Barriers to Technology Adoption When Fostering Entrepreneurship Among the Poor: The Role of Technology and Digital Literacy. *IEEE Transactions on Engineering Management*, *68*(6), 1605–1618.

Peeters, M. J., & Augustine, J. M. (2023). Using Rasch measurement for instrument rating scale refinement. *Currents in Pharmacy Teaching and Learning*, *15*(1), 110–118.

Peeters, M. J., & Harpe, S. E. (2022). Contemporary conceptualization of measurement validity. In *Contemporary Research Methods in Pharmacy and Health Services* (pp. 575–582). Elsevier.

Pfeifer, K., Ergal, İ., Koller, M., Basen, M., Schuster, B., & Rittmann, S. K.-M. R. (2021). Archaea Biotechnology. *Biotechnology Advances*, *47*, 107668.

Planinic, M., Boone, W. J., Susac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research: Why measurement matters. *Physical Review Physics Education Research*, *15*(2), 20111.

Purniasari, L., Masykuri, M., & Ariani, S. R. D. (2021). Analisis Butir Soal Ujian Sekolah Mata Pelajaran Kimia Sma N 1 Kutowinangun Tahun Pelajaran 2019/2020 Menggunakan Model Iteman Dan Rasch. *Jurnal Pendidikan Kimia*, *10*(2), 205–214.

Putranta, H., Jumadi., J., & Wilujeng, I. (2019). Physics learning by PhET simulation-assisted using problem based learning (PBL) model to improve students' critical thinking skills in work and energy chapters in MAN 3 Sleman. In *Asia-Pacific Forum on Science Learning and Teaching* (Vol. 20, No. 1, pp. 1-44).

Raharjo, S., Kaniawati, I., & Suwarma, I. R. (2014). *Technology and Engineering Literacy Profile of Senior High School Students ' on Understanding The Newton Laws '*. 3–8.

Rahmawati, T., Mulyaningsih, T., Nahadi, N., Suhanda, H., Lee, W. K., Aziz, H. A., & Anwar, S. (2023). Electronic Portfolio Assessment Instruments in Improving Students' Creative Thinking Skills. *Jurnal Pendidikan IPA Indonesia*, *12*(4), 598–610.

Schaufeli, W. B., Desart, S., & De Witte, H. (2020). Burnout Assessment Tool (BAT)—Development, Validity, and Reliability. *International*

Journal of Environmental Research and Public Health, *17*(24), 9495.

Silvia, P. J., Rodriguez, R. M., Beaty, R. E., Frith, E., Kaufman, J. C., Loprinzi, P., & Reiter-Palmon, R. (2021). Measuring everyday creativity: A Rasch model analysis of the Biographical Inventory of Creative Behaviors (BICB) scale. *Thinking Skills and Creativity*, *39*, 100797.

Sudaryono, S., Margono, G., & Rahayu, W. (2013). *Pengembangan Instrumen Penelitian Pendidikan*. Yogyakarta: Graha Ilmu.

Sumintono, B., & Widhiarso, W. (2014). *Aplikasi model Rasch untuk penelitian ilmu-ilmu sosial (edisi revisi)*. Trim Komunikata Publishing House.

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan RASCH Pada Assessment Pendidikan* (C. T. Komunikata (ed.)).

Supriyadi, T., Saptani, E., Rukmana, A., Suherman, A., Alif, M. N., & Rahminawati, N. (2020). Students' technological literacy to improve academic writing and publication quality. *Universal Journal of Educational Research*, *8*(11B), 6022–6035.

Susanta, A., & Susanto, E. (2021). The development of test instruments to support students' ability on mathematical learning evaluation course. *Journal of Physics: Conference Series*, *1731*(1).

Taher, M., & Abtaria, Y. (2017). Efektifitas Pembelajaran Creative Problem Solving Berbasis Eksperimen dalam Pembelajaran Fisika untuk Melatihkan Literasi Teknologi Siswa. *Gravity*, *3*(2), 148–157.

Triana, A. J., Gusdorf, R. E., Shah, K. P., & Horst, S. N. (2020). Technology Literacy as a Barrier to Telehealth During COVID-19. *Telemedicine and E-Health*, *26*(9), 1118–1119.

Upadhayaya, P. R. (2023). Information Communication Technology in Education: Bringing Innovation in Classroom. *Ganeshman Darpan*, *8*(1), 96–110.

Wensveen, J. (2023). *Air Transportation: A Global Management Perspective*. Routledge.

Widoyoko, E. P. (2012). *Teknik Penyusunan Instrumen Penelitian*. Yogyakarta: Pustaka Pelajar.

Willis, E., Gundacker, C., Harris, M., & Mameledzija, M. (2019). Improving immunization and health literacy through a community-based approach enhanced by technology. *Information Services and Use*, *39*(1–2), 23–36.

Yan, Q., Dong, H., Su, J., Han, J., Song, B., Wei, Q., & Shi, Y. (2018). A Review of 3D Printing Technology for Medical Applications. *Engineering*, *4*(5), 729–742.

Yu-Te, W., Yuan-Tai, C., Shun, L. Y., & Kai-Fu, Y. (2014). A survey of Energy Technology literacy level for senior high school students in Taiwan. In *Proceedings of International Conference on Engineering and Computer Education* (Vol. 6).

Zainurrisalah, T. F., Suwarma, I. R., & Jauhari, A. (2018). Mengukur Kemampuan Literasi Teknologi dan Rekayasa (Engineering) Melalui Penerapan Pembelajaran STEM dalam Fisika. *Prosiding Seminar Nasional Fisika (E-Journal)*, 131–135.