24 (2), 2023, 419-431

Jurnal Pendidikan MIPA

e-ISSN: 2685-5488 | p-ISSN: 1411-2531

http://jurnal.fkip.unila.ac.id/index.php/jpmipa/

Using the Rasch Model to Develop a Measure of Students' Problem Solving Ability in Optical Instruments

Rosi Vera*, Ida Kaniawati, & Judhistira Aria Utama

Department of Physics Education, Indonesia University of Education, Indonesia

Abstract: This study aims to develop a test instrument for problem-solving abilities in the material of optical instruments. The research method used is the development of a 4D model consisting of defining, designing, developing, and disseminating. Participants in this study consisted of 75 consisting of 40 males and 35 females in grade 12 high school sciences program at a high school in Bandung. Based on the results of the analysis, it was obtained: (1) expert validation with a CVR index value of 1 and the test instrument included in the valid category, (2) construct validation using Rasch was declared valid, (3) the reliability of each item using Rasch analysis was declared reliable with Cronbach alpha of 0.96, and (4) the difficulty level of the items is evenly distributed at each level of difficulty. Therefore, the test instrument that has been developed can measure students' problem-solving abilities in the material of optical instruments.

Keywords: pemodelan Rasch, problem solving ability, optical instruments.

Abstrak: *Penelitian ini bertujuan untuk mengembangkan intrumen test kemampuan pemecahan masalah pada materi alat-alat optik . Metode penelitian yang digunakan yaitu pengembangan model 4D yang terdiri dari defining, designing, developing, dan disseminating. Partisipan dalam penelitian ini terdiri dari 75 yang terdiri dari 40 laki-laki dan 35 perempuan pada kelas 12 SMA program Ilmu Pengetahuan Alam di salah satu SMA di Bandung. Berdasarkan hasil analisis diperoleh: (1) validasi ahli dengan nilai CVR indeks sebesar 1 dan instrumen tes termasuk dalam kategori valid, (2) validasi kontruk menggunakan Rasch dinyatakan valid, (3) reabilitas setiap butir soal dengan analisis Rasch dinyatakan reliabel dengan Cronbarch Alpha sebesar 0,96, dan (4) tingkat kesukaran butir soal merata pada setiap tingkat kesukaran. Oleh karena itu, instrumen tes yang telah dikembangkan dapat mengukur kemampuan pemecahan masalah peserta didik pada materi alat-alat optik.*

Kata kunci: *pemodelan Rasch, kemampuan pemecahan masalah, instrument optik.*

▪ INTRODUCTION

Problem solving ability is one of the most important skills in 21st century education (Ince, 2018; Tang, Vezzani & Erikson, 2020). Problem solving ability as a cognitive ability that is used to solve problems related to the real world (OECD, 2003). This relates to individual skills in finding a solution to a problem in everyday life by involving knowledge (Bahtiyar & Can, 2016). This is intended to prepare students to be able to solve not only structured but also unstructured, complex, and diverse problems (Dixon & Brown, 2012). Problem solving ability is how one thinks or looks for ways to reason, in applying the knowledge that has been acquired to overcome the problems encountered so as to get a solution to a problem (Heller & Heller, 2010; Apriyani et al., 2019).

It is very important to apply problem-solving ability in the physics learning process to help students deepen their understanding of a scientific concept in order to

Rosi Vera
*Email: rosivera@upi.edu

DOI: http://dx.doi.org/10.23960/jpmipa/v24i2.pp419-431
Received: 18 April 2022
Accepted: 12 June 2022
Published: 20 June 2022

systematically analyze information in making critical reflections (Scherer & Beckmann, 2014; Xie & Mintaila, 2017; Lim & Han, 2020). The ability to solve problems in physics can guide students to solve complex problems not only in class but also in the world of work (Williams, 2018; Purwaningsih et al., 2020). Learning physics contains questions of everyday life, problem solving skills are very important in learning physics because it is not enough for students to see only the quantitative aspects of solving these equations and how mathematics is used, but also qualitative analysis to choose the concepts and principles correctly in answer questions (Docktor et al., 2015). In learning physics, students' cognitive abilities are very helpful in solving physics problems, such as skills to support and understand physics principles correctly, ability to understand and understand physics problems effectively (Permatasari et al., 2018).

In accordance with the Regulation of the Minister of Education and Culture No. 21 of 2016 concerning Process Standards for Elementary and Secondary Education states that the learning process carried out in schools is expected to be able to train students' problem-solving abilities (Kemendikbud, 2016). Problem solving ability is associated with critical thinking, analytical and productive creation which involves quantitative, communicative, and the ability to respond critically (Chang, 2010). This shows that it is very important to develop a problem-solving ability test instrument to measure students' problem-solving abilities.

The large number of studies on problem solving abilities certainly requires test instruments that are used to measure problem solving abilities (Gok & Silay, 2008; Waler & Kaye, 2012; Aristiawan & Istiyono, 2020). Most of the problem solving ability test instruments are multiple choice (J. Sirait et al., 2017; Istiyono et al., 2019; Suprapto et al., 2020). However, multiple-choice tests have a drawback where students expect answers that are available in the answer choices (Kubiszyn & Borich, 2016). In addition, multiple choice tests do not provide complete information on the problem solving abilities measured in students (Kastner & Stangl, 2011). Another factor, measurement results are less accurate because many students are able to answer questions correctly but do not know the relationship between concepts and questions (Henderson et al., 2001). Therefore, another type of test is needed to measure the level of students' problem-solving abilities.

Essay test is an instrument that can be used to measure problem-solving skills (Haladyna & Rodriguez, 2013; Aristiawan & Istiyono, 2020). The essay test instrument is able to reveal the thought process in solving problems (Winarto et al., 2022). Essay tests can measure complex and comprehensive cognitive abilities in solving problems (Haladyna & Rodriguez, 2013). In addition, essay tests can encourage students to think at a higher level in answering questions (Baig et al., 2014). Therefore, essay tests are suitable for measuring students' problem-solving abilities.

This study developed an essay test instrument based on indicators developed by (Docktor et al., 2016). The indicators of problem solving ability are as follows: namely 1) usefull description, 2) physics approach, 3) specific application of physics, 4) mathematical procedure, and 5) logical progression. The test instrument developed in this study focuses on the material of optical instruments. The selection of this material is based on the fact that there are students' difficulties in the material of optical instruments (Rokhmah, Sunarno & Masykuri 2017; Tumanggor et al, 2018). In addition, other research shows that students have difficulty understanding the concepts of light

and vision in physics (Uzun, Alev & Karal 2013). Other research states that students still have difficulty understanding the material for forming images in mirrors (Permatasari et al., 2018).

▪ METHOD

Participant

Participants in this study involved 75 students consisting of 40 boys and 35 girls in grade 12 high school sciences program. This research was conducted in the even semester of the 2022/2023 academic year. Students were selected using random sampling. Selection of samples random sampling is done randomly so that each member of the population has the same opportunity to be selected.

Research Design Procedures

The research method used in this research is the 4D instructional research and development method (define, design, develop, and disseminate). The development procedure in this study followed the 4D method as follows: 1) defining the problem-solving ability test instrument, 2) designing the problem-solving ability test instrument, 3) developing the problem-solving ability test instrument, and 4) disseminating the problem-solving ability test instrument.

In the define stage, a literature review and analysis was carried out on the 2013 revision of the 2019 curriculum to determine students' problem-solving abilities in the material of optical devices. Furthermore, at the design stage, it examines the suitability of the items with indicators, answer keys, scoring guidelines, and the suitability of the concepts used. At the develop stage, the instrument is arranged according to the initial design. Furthermore, expert validation and empirical validation were carried out using the rash model. Expert validation aims to obtain an expert's assessment of problem solving abilities. In addition, expert validation also aims to obtain suggestions for improving the instrument. If the instrument is declared feasible by the expert, an empirical test is then carried out to find out the validity and reliability of the instrument's ability to solve problems in the material of optical instruments. The disseminate stage is used to test the instrument on students.

Instruments

The instrument used in this study is a problem solving ability test instrument. The test instrument consists of six essay questions on optical instruments. Each item developed in this test refers to the problem-solving ability indicators proposed by Docktor et al (2016), namely 1) useful description, 2) physics approach, 3) specific application of physics, 4) mathematical procedure, and 5) logical progression. The distribution of the item indicators is presented in Table 1.

Table 1. Distribution of item indicators

Question indicators	Item Number
Comparing the eyes of hypermetropic sufferers to the location of an object	1
Evaluate lens strength in people with eye defects	2
Analyze the use of lenses and magnification on the microscope	3

Analyze the size of objects using a microscope	4
Analyze the use of lenses and magnification in telescopes	5
Analyze the use of the telescope from minimum accommodation to accommodation at a certain distance by designing the position of the ocular lens to the objective lens on the telescope	6

Data Analysis

In this study, the instrument for testing students' problem-solving abilities was analyzed through four stages. The first stage is the analysis of the results of the expert validation using the Content Validity Ratio (CVR) and the Rasch multi-rater measurement. The CVR value of the expert validation results on the problem solving ability test instrument is calculated using equation 1.

$$\mathrm{CVR} = \frac{\mathrm{n_e} - \frac{\mathrm{N}}{2}}{\frac{\mathrm{N}}{2}}$$

Information :
n_e : Total score for each aspect of the assessment
N : Total number of validators

The results of calculating the CVR value are interpreted according to what is presented in Table 2 below:

Table 2. Nilai CVR minimum value α = 0.5 (Wilson et al., 2012)

Number of Validators	$CVR_{critical}$
5	0.736
6	0.672
7	0.622
8	0.582
9	0.548
10	0.520

In this study, the instrument is included in the valid category if the index CVR ≥ critical CVR. However, the instrument is included in the invalid category if the CVR index <critical CVR. The critical CVR for the number of validators consisting of five validators is 0.736 (Wilson et al., 2012). The second stage determines construct validity by using Rasch analysis. Construct validity is determined by the Outfit mean square (MNSQ), Outfit Z-standard (ZSTD), and Measure Correlation (Pt Mean Corr) (Sumintono & Widhiarsho, 2015). The details of the three criteria are shown in Table 3.

Table 3. Details of the question validity criteria

Outfit	Accepted ranges
Outfit mean square (MNSQ)	0.5 < MNSQ < 1.5
Outfit Z-standard (ZSTD)	-2.00 < ZSTD < +2.00
Point Measure Correlation (Pt.Mean Corr)	0.4 < Pt. M. corr < 0.85

Construct validity is also carried out by a unidimensionality test which aims to ensure that the instruments used are appropriate and can be used to measure the variables studied. This test was carried out by analyzing the values of raw variance explained by measures and unexplained variance 1st contrast (Nurdini et al., 2020). The third stage is the reliability test using Rasch analysis. The reliability test in this study aims to show consistent results of measurement of problem-solving ability test instruments (Sumintono & Widhiarso, 2015). The instrument is included in the reliable category if it meets the Cronbach alpha, person reliability and item reliability values. Interpretation of Cronbach alpha values is shown in Table 3. Interpretation of person reliability and item reliability values is shown in Table 4.

Table 4. Criteria of Cronbach alpha value (Sumintono & Widhiarso, 2015)

Range value	Category
$0.00 \leq r < 0.50$	Very bad
$0.50 \leq r < 0.60$	Bad
$0.60 \leq r < 0.70$	Enough
$0.70 \leq r < 0.80$	Good
$0.80 \leq r < 1.00$	Very good

Table 5. Criteria for the value of person reliability and item reliability (Sumintono & Widhiarso, 2015)

Person reliability and item reliability	Category
$0.00 \leq r < 0.67$	Weak
$0.67 \leq r < 0.81$	Enough
$0.81 \leq r < 0.91$	Good
$0.91 \leq r < 0.94$	Very good
$0.94 \leq r < 1.00$	Special

The fourth stage determines the level of difficulty of each item analyzed using Rasch. Difficulty level analysis aims to find out which test instruments that have been prepared are included in the easy or difficult category. The analysis of the difficulty level of the items is reviewed based on the measure value in logit units (Sumintono & Widhiarso, 2015). The interpretation of the difficulty level of the questions is shown in the following table:

Table 6. Interpretation of the difficulty level of the items

Logit Measured Value	Interpretation
Very easy	$M < -1SD$
Easy	$-1SD \leq M \leq 0$
Hard	$0 \leq M \leq SD$
Very Difficult	$M > SD$

▪ RESULT AND DISSCUSSION

The results and discussion in the research related to the development of problem solving ability test instruments using 4D (define, design, develop, and disseminate) will be discussed as follows:

Define

At the defining stage in this research is to conduct a literature review related to students' problem-solving abilities and carry out an analysis of the 2013 revised 2019 curriculum related to the material of optical devices. Based on this, students' problem solving abilities can be measured through the indicators put forward by Docktor et al (2016), namely 1) useful description, 2) physics approach, 3) specific application of physics, 4) mathematical procedure, and 5) logical progression.

Design

Stages of test instrument design are made based on indicators and rubrics for assessing problem solving abilities. The sub-matter of the problem-solving ability test on optical instruments consists of: 1) eyes and glasses; 2) microscope and 3) binoculars. This test instrument is in the form of description questions which total 6 questions. The distribution of each question based on the problem-solving ability indicator for each question can be seen in Table 1 below:

Table 7. Distribution of Problem Solving Ability Test Questions

No	Indicator of problem solving ability	Number of Questions
1.	useful description	1a. 2a. 3a. 4a. 5a dan 6a
2.	physics approach	1b. 2b. 3b. 4b. 5b. dan 6b
3.	specific application of physics	1c. 2c. 3c. 4c. 5c. dan 6c
4.	mathematical procedure	1d. 2d. 3d. 4d. 5d. dan 6d
5.	logical progression	1e. 2e. 3e. 4e. 5e dan 6e

Develop

At the develop stage, each question item is made based on an indicator of problem solving ability which consists of six essay questions. Figure 1 shows one of the questions on the problem solving ability test instrument. After all the items on the problem solving ability test are made. Furthermore, judgment was made by five experts on the problem-solving ability test instrument that had been developed. Five experts were selected to provide judgment. The assessment aspects that are assessed are material, construction, and language. The expert validation sheet used consisted of three criteria namely "valid without revision", "valid with revision", and "invalid". Then the results of the expert assessment were analyzed using the Content Validity Ratio (CVR). To get valid or invalid conclusions for each item that has been validated by five experts (N = 5) with a critical CVR of 0.736. Based on the results of expert validation of the item problem solving ability test instrument, the CVR index value is 1. This indicates that the CVR index value > CVR is critical. In line with research conducted by (Yudhistira & Tomoliyus, 2020; Hidayat et al., 2022) states that a CVR index close to 1 indicates good or high validity. In line with another study conducted by (Sundari et al., 2023) stated that a CVR index value of more than 0 indicates that the item is valid. Furthermore, the results of expert validation were analyzed using the multi-rater Rasch measurement test. The results of expert assessments analyzed using a multi-rater are shown in Figure 2.

No	Question Indicator	Question Items	Question Answers	Score
1	Comparing the eyes of hypermetropic sufferers to the location of an object	Mrs. Ulfa and Mrs. Cut are biology teachers and math teachers at a school. Ulfa's mother has a visual impairment which causes her to be able to see distant objects clearly and a near point of 100 cm. Mrs. Cut also has an eye disorder with a near point of 50 cm. They use glasses to help correct vision at a normal distance (25 cm). Once upon a time, the two of them were in a hurry to enter the classroom because they weren't careful, their glasses were switched and their eyesight became different from before. You will solve the problem regarding the distance that Ms. Ulfa and Ms. Cut can see clearly when wearing the exchanged glasses. a. Describe the magnitude and units related to the visual impairment experienced by Mrs. Ulfa and Mrs. Cut! b. Write down the physics approach as a solution to calculate the distance they can see clearly when wearing the glasses swapped! c. Write down the mathematical equation used to calculate the distance they can see clearly when wearing the glasses that are swapped! d. Calculate the results of solving the problem using a mathematical equation! e. Write a conclusion from the solution to solving the problem!	***Useful Description*** : Given: $s'_{Mrs\ Ulfa} = 100$ cm ; $s'_{Mrs\ Cut} : 50$ cm ; $s_n : 25$ cm Mrs. Ulfa wears Mrs. Cut's glasses, and vice versa. Asked: How far can they see clearly when wearing interchanged glasses ($s_{Mrs\ Ulfa}$ dan $s_{Mrs\ Cut}$) ?	5
			Physics Approach: Mrs. Ulfa and Mrs. Cut both suffer from nearsightedness (hypermetropia) because they have a near point of more than 25 cm and a far point at infinity. Therefore, people who suffer from nearsightedness can clearly see objects that are very far away without accommodation, but cannot see close objects clearly.	5
			Specific Aplication of Physics : Hypermetropia eye defects can be overcome by using convex (positive) lens glasses. In order to see clearly the image produced must be virtual s' = -. The equation used in patients with hypermetropia is as follows: $\frac{1}{f} = \frac{1}{s} + \frac{1}{s'}$ become $\frac{1}{f} = \frac{1}{s_n} + \frac{1}{s'}$	5
			Math Procedures : • Looking for Mrs. Ulfa's focus point in seeing $\frac{1}{f_{Mrs.Ulfa}} = \frac{1}{s_n} + \frac{1}{s'_{Mrs.Ulfa}}$ $\frac{1}{f_{Mrs.Ulfa}} = \frac{1}{25} + \frac{1}{-100\ cm}$ $\frac{1}{f_{Mrs.Ulfa}} = \frac{4-1}{100\ cm} = \frac{3}{100}\ cm$ $f_{Mrs.Ulfa} = 33{,}33\ cm$ So, the focal point of Mrs. Ulfa to see normally is 33.33 cm • Looking for a focal point Mrs. Cut in view $\frac{1}{f_{Mrs.Cut}} = \frac{1}{s_n} + \frac{1}{s'_{Mrs.Cut}}$	5
			$\frac{1}{f_{Mrs.Cut}} = \frac{1}{25} + \frac{1}{-50\ cm}$ $\frac{1}{f_{Mrs.Cut}} = \frac{2-1}{50\ cm} = \frac{1}{50}\ cm$ $f_{Mrs.Cut} = 50\ cm$ So, the focal point of Mrs. Cut to see normally is 50 cm • To find out the far point so that Mrs. Ulfa and Mrs. Cut can see clearly is as follows: $\frac{1}{s_{Mrs.Ulfa}} = \frac{1}{f_{Mrs.Cut}} + \frac{1}{s'_{Mrs.Ulfa}}$ $\frac{1}{s_{Mrs.Ulfa}} = \frac{1}{50}\ cm + \frac{1}{100}\ cm$ $\frac{1}{s_{Mrs.Ulfa}} = \frac{2+1}{100}\ cm = \frac{3}{100}\ cm$ $s_{Mrs.Ulfa} = 33{,}33\ cm$ So, the point near Mrs. Ulfa in sight is 33.33 cm $\frac{1}{s_{Mrs.Cut}} = \frac{1}{f_{Mrs.Ulfa}} + \frac{1}{s'_{Dira}}$ $\frac{1}{s_{Mrs.Cut}} = \frac{3}{100}\ cm + \frac{1}{50}\ cm$ $\frac{1}{s_{Mrs.Cut}} = \frac{3+2}{100} = \frac{5}{100}$ $s_{Mrs.Cut} = 20\ cm$ So, the point near Mrs. Cut in saw is 20 cm	
			Logical Progression: Based on the problems above, it can be concluded that Mrs. Ulfa and Mrs. Cut both suffer from hypermetropic eye defects. Patients with hypermetropia can be overcome by using convex (positive) lens glasses. Thus, Mrs. Ulfa can see clearly when wearing Mrs. Cut at t near point 33, 33 cm while Mrs. Cut while using Mrs. glasses. Ulfa is at a point near 20 cm.	5

Figure 1. Problem solving ability test instrument items

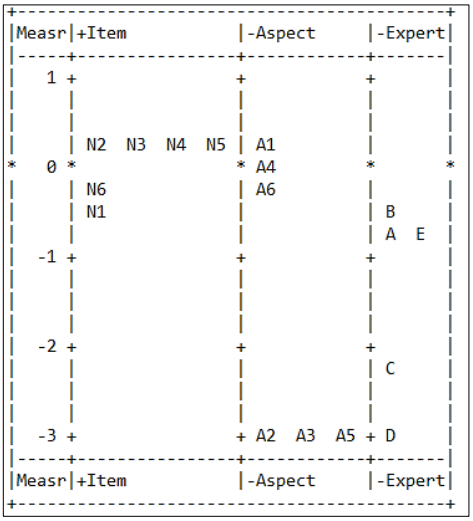


Figure 2. Multi-rater results of expert validation of problem-solving ability test instruments

Based on Figure 2, the results of the multi-rater analysis by five experts (A, B, C, D and E) show that the problem-solving ability items that meet the requirements are N2, N3, N4, N5. Whereas the items N6 and N1 were considered not good according to the expert. The difficulty level of the item is shown in the third column, the higher the logit value of an item assessment indicator, the more difficult the item assessment indicator is met by the validator (expert). Conversely, the lower the logit value of an item assessment indicator, the easier it is to fulfill the assessment indicator (Darmana et al., 2021). Meanwhile, the assessment aspects that were the most difficult to achieve were A1, A4 and A6. According to the validator, the aspects of the assessment that are easiest to fulfill are A2, A3 and A5.

The assessment of each validator can also be analyzed to determine the reliability of the validator. Based on the analysis using Rasch it is known that Experts B and A are the most consistent with the validity criteria. This can be seen from the Outfit Mean Square (MNSQ) values obtained by each expert A and B which are 1.15 and 1.04. These results indicate that the Outfit MNSQ criteria are met because they are in the range 0.5 < MNSQ < 1.5 (Boone et al., 2014; Darmawan et al., 2021). Furthermore, in terms of the Outfit Z-standard (ZSTD) values of experts A and B, they are 1.0 and 0.2 respectively. These results meet the Outfit ZSTD criteria because they are in the range -2.00 < ZSTD < +2.0 (Boone et al., 2014; Darmawan et al., 2021).

Whereas Experts E and C showed an invalid assessment because they did not meet the Outfit ZSTD score criteria. Whereas expert D is the expert who has the lowest Outfit MNSQ and Outfit ZSTD scores. Based on expert judgment, a reliability value of 0.51 was obtained. This shows that the experts give quite different scores, but some are the same (Kocak, 2020). The similarity of judgments by five experts (correct agreement) is 78.9%. This shows that there is a slight difference in the assessment of each item on the problem solving ability test. This is in line with research conducted by (Brookhart et al., 2006; Güler, 2014; Darmana et al., 2021) which states that the tendency for assessment is different because the assessor's behavior such as leniency and severity affects the assessor's condition.

Disseminate

The disseminate stage of the questions that had been developed was tested on 75 students of class XII in the Natural Sciences program at a public high school in Bandung. Furthermore, an analysis was carried out to determine the validity of the items. The following are the results of the analysis for each item on the problem-solving ability test as shown in Figure 3:

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE 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
13	162	75	.22	.10	1.95	4.66	1.92	4.14	A .52	.72	14.7	33.4	N3(c)
18	177	75	.07	.10	1.76	4.07	1.88	4.12	B .52	.70	20.0	31.6	N4(c)
28	137	75	.48	.11	1.70	3.31	1.49	2.26	C .65	.75	21.3	35.8	N6(c)
10	220	75	-.32	.09	1.59	3.52	1.40	2.17	D .55	.63	22.7	31.1	N2(e)
8	162	75	.22	.10	1.54	2.90	1.46	2.33	E .56	.72	33.3	33.4	N2(c)
23	192	75	-.07	.10	1.18	1.19	1.40	2.15	F .52	.68	26.7	30.8	N5(c)
27	204	75	-.18	.09	1.13	.88	1.26	1.51	G .57	.66	20.0	31.8	N6(b)
19	173	75	.11	.10	1.25	1.55	1.25	1.38	H .60	.71	21.3	32.0	N4(d)
9	202	75	-.16	.09	1.08	.60	1.07	.47	I .62	.66	36.0	31.6	N2(d)
11	169	75	.15	.10	1.03	.22	1.06	.41	J .73	.71	32.0	32.4	N3(a)
17	154	75	.30	.10	1.05	.34	1.06	.37	K .71	.73	30.7	33.9	N4(b)
5	226	75	-.38	.09	1.03	.22	.97	-.14	L .64	.62	28.0	31.3	N1(e)
30	226	75	-.38	.09	1.02	.20	.92	-.40	M .63	.62	21.3	31.3	N6(e)
26	168	75	.16	.10	.97	-.11	1.00	.04	N .79	.72	14.7	32.6	N6(a)
24	133	75	.53	.11	.90	-.49	.94	-.27	O .83	.76	24.0	36.6	N5(d)
1	224	75	-.36	.09	.91	-.59	.82	-1.07	o .68	.63	25.3	31.3	N1(a)
6	220	75	-.32	.09	.75	-1.81	.91	-.51	n .68	.63	32.0	31.1	N2(a)
15	182	75	.02	.10	.90	-.66	.85	-.85	m .74	.70	34.7	31.0	N3(e)
21	193	75	-.08	.10	.90	-.66	.87	-.76	l .71	.68	29.3	31.0	N5(a)
22	243	75	-.53	.10	.87	-.89	.90	-.54	k .61	.59	40.0	34.9	N5(b)
4	205	75	-.19	.09	.89	-.72	.77	-1.42	j .73	.66	44.0	31.7	N1(d)
14	191	75	-.06	.10	.86	-.94	.82	-1.07	i .71	.68	36.0	30.8	N3(d)
16	205	75	-.19	.09	.82	-1.27	.85	-.86	h .67	.66	36.0	31.7	N4(a)
25	149	75	.35	.10	.78	-1.32	.83	-.90	g .82	.74	28.0	34.4	N5(e)
12	144	75	.41	.10	.79	-1.26	.71	-1.62	f .83	.75	40.0	34.5	N3(b)
3	193	75	-.08	.10	.71	-2.13	.75	-1.59	e .75	.68	36.0	31.0	N1(c)
7	210	75	-.23	.09	.70	-2.31	.73	-1.69	d .74	.65	29.3	30.6	N2(b)
29	160	75	.24	.10	.73	-1.80	.65	-2.18	c .84	.73	44.0	33.3	N6(d)
20	142	75	.43	.11	.72	-1.72	.69	-1.75	b .82	.75	45.3	34.8	N4(e)
2	202	75	-.16	.09	.55	-3.71	.55	-3.21	a .82	.66	30.7	31.6	N1(b)
MEAN	185.6	75.0	.00	.10	1.03	.04	1.03	.02			29.9	32.4	
P.SD	29.7	.0	.28	.00	.34	1.97	.34	1.76			8.2	1.6	

Figure 3. Outfit (mnsq), outfit zstd, and pt measure core scores for each problem-solving ability test item

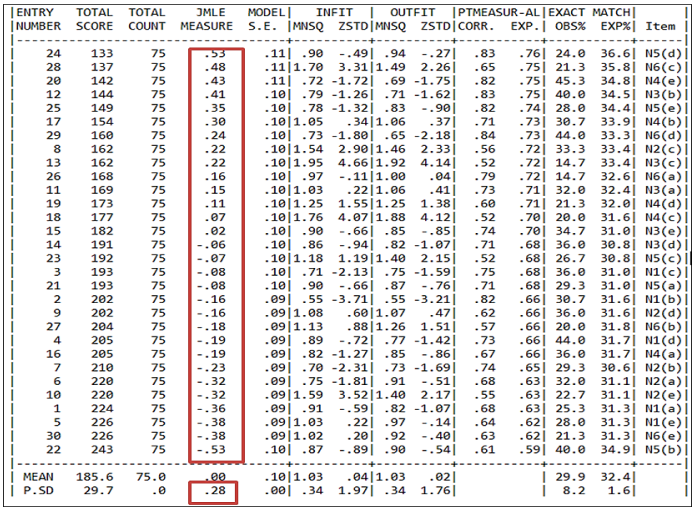
Based on Figure 3, the results of the analysis of problem solving ability items using Rasch obtained Outfit MNSQ values in the range of 0.55 – 1.49. This indicates that the MNSQ Outfit value is fulfilled because it is in the range 0.5 < MNSQ < 1.5 (Boone et al., 2014; Darmawan et al., 2021). However, the MNSQ Outfit scores for items N3(c) and N4(c) are outside the specified criteria range. Judging from the Outfit ZSTD value on each test item problem solving ability is obtained in the range -1.75 – 1.51. This shows that the Outfit ZSTD value is fulfilled because it is in the range -2.00 < ZSTD < +2.0. However, in item N3(c), N4(c), N6(c), N2(c), N2(e), N5(c), N1(b) and N6(d) the ZSTD Outfit value is outside the criteria which is determined. The PT Mean Corr value for each item is in the range of 0.52-0.82. this shows that the PT Mean Corr obtained in each item meets the criteria because it is in the range 0.4 < PT Mean Corr < 0.85 (Sumintono & Widhiarso, 2015). This shows that the items on the problem solving ability test can be said to be appropriate. The quality of the item items can be used if only one or two criteria are met, then these items can be maintained and said to be "appropriate" without needing to be changed (Dewi et al., 2021; Nurdini et al., 2020).

Meanwhile, if all of these criteria are not met, then the item items can be said to be inappropriate and the item items need to be repaired or replaced (Dewi et al., 2021; Nurdini et al., 2020). Therefore, it can be concluded that all items on the problem solving ability test are valid and can be used without needing to be repaired.

The instrument validation test is also reviewed through the instrument's unidimensionality test which aims to ensure that the instrument is feasible to use to measure the variables studied (Nurdini et al., 2020). This test was carried out by analyzing the value of raw variance explained by measures and unexplained variance 1st contrast (Nurdini et al., 2020). The requirements for the unidimensionality test are that the eigen value on the unexplained variance 1st contrast has a value of less than 3, while the observed value has a value of less than 15% (Mofreh et al., 2014; Ng et al., 2018; Sumintono & Widhiarso, 2015; Talib et al., 2019). Based on the results of the analysis using Rasch, the raw variance explained by measures is 51%. This shows that the item is included in the good category. This is in line with research conducted by (Ng et al., 2018; Talib et al., 2019) which states that more than 40% of polytomous data is included in a good interpretation. Furthermore, the unexplained variance 1st contrast on the observed value obtained a value of 6.2%, so the item meets the criteria because the observed value is less than 15%. In addition, the unexplained variance 1st contrast on the eigen value obtained a value of 2.77. This shows that the items meet the criteria because the value is less than which indicates that the instrument does not measure only one dimension (Talib et al., 2019). Thus, the results of the unidimensionality test are valid for measuring students' problem-solving abilities.

In addition to validity analysis, reliability analysis was also carried out on the problem-solving ability test items. The reliability test aims to show consistent measurement results of the test instrument (Sumintono & Widhiarso 2015). Based on the results of the reliability test using Rasch analysis, the person reliability value was 0.95. This value indicates that the consistency of students in answering test instruments is included in the special category. Furthermore, the item reliability value obtained is 0.86 so that the problem solving ability test instrument has reliability in the good category. Meanwhile, the Cronbach Alpha value was 0.96 so that the agreement between items and persons (students) was included in the special category. Therefore, it can be concluded that the problem solving ability test instrument compiled is reliable (Sumintono & Widhiarso 2015).

Next, an analysis of the difficulty level of the problem-solving ability test items was carried out using rasch. This analysis aims to find out the test instruments that have been compiled are included in the easy or difficult categories. According to Sumintono & Widhiarso (2015) the output of Table 13 Item Measure can be used to determine the level of difficulty of the items by using the Standard Deviation (SD) value and logit value. The output results of the item measure the difficulty level of the problem-solving ability test items as shown in the following figure:



ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE 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
24	133	75	.53	.11	.90	-.49	.94	-.27	.83	.76	24.0	36.6	N5(d)
28	137	75	.48	.11	1.70	3.31	1.49	2.26	.65	.75	21.3	35.8	N6(c)
20	142	75	.43	.11	.72	-1.72	.69	-1.75	.82	.75	45.3	34.8	N4(e)
12	144	75	.41	.10	.79	-1.26	.71	-1.62	.83	.75	40.0	34.5	N3(b)
25	149	75	.35	.10	.78	-1.32	.83	-.90	.82	.74	28.0	34.4	N5(e)
17	154	75	.30	.10	1.05	.34	1.06	.37	.71	.73	30.7	33.9	N4(b)
29	160	75	.24	.10	.73	-1.80	.65	-2.18	.84	.73	44.0	33.3	N6(d)
8	162	75	.22	.10	1.54	2.90	1.46	2.33	.56	.72	33.3	33.4	N2(c)
13	162	75	.22	.10	1.95	4.66	1.92	4.14	.52	.72	14.7	33.4	N3(c)
26	168	75	.16	.10	.97	-.11	1.00	.04	.79	.72	14.7	32.6	N6(a)
11	169	75	.15	.10	1.03	.22	1.06	.41	.73	.71	32.0	32.4	N3(a)
19	173	75	.11	.10	1.25	1.55	1.25	1.38	.60	.71	21.3	32.0	N4(d)
18	177	75	.07	.10	1.76	4.07	1.88	4.12	.52	.70	20.0	31.6	N4(c)
15	182	75	.02	.10	.90	-.66	.85	-.85	.74	.70	34.7	31.0	N3(e)
14	191	75	-.06	.10	.86	-.94	.82	-1.07	.71	.68	36.0	30.8	N3(d)
23	192	75	-.07	.10	1.18	1.19	1.40	2.15	.52	.68	26.7	30.8	N5(c)
3	193	75	-.08	.10	.71	-2.13	.75	-1.59	.75	.68	36.0	31.0	N1(c)
21	193	75	-.08	.10	.90	-.66	.87	-.76	.71	.68	29.3	31.0	N5(a)
2	202	75	-.16	.09	.55	-3.71	.55	-3.21	.82	.66	30.7	31.6	N1(b)
9	202	75	-.16	.09	1.08	.60	1.07	.47	.62	.66	36.0	31.6	N2(d)
27	204	75	-.18	.09	1.13	.88	1.26	1.51	.57	.66	20.0	31.8	N6(b)
4	205	75	-.19	.09	.89	-.72	.77	-1.42	.73	.66	44.0	31.7	N1(d)
16	205	75	-.19	.09	.82	-1.27	.85	-.86	.67	.66	36.0	31.7	N4(a)
7	210	75	-.23	.09	.70	-2.31	.73	-1.69	.74	.65	29.3	30.6	N2(b)
6	220	75	-.32	.09	.75	-1.81	.91	-.51	.68	.63	32.0	31.1	N2(a)
10	220	75	-.32	.09	1.59	3.52	1.40	2.17	.55	.63	22.7	31.1	N2(e)
1	224	75	-.36	.09	.91	-.59	.82	-1.07	.68	.63	25.3	31.3	N1(a)
5	226	75	-.38	.09	1.03	.22	.97	-.14	.64	.62	28.0	31.3	N1(e)
30	226	75	-.38	.09	1.02	.20	.92	-.40	.63	.62	21.3	31.3	N6(e)
22	243	75	-.53	.10	.87	-.89	.90	-.54	.61	.59	40.0	34.9	N5(b)
MEAN	185.6	75.0	.00	.10	1.03	.04	1.03	.02			29.9	32.4	
P.SD	29.7	.0	.28	.00	.34	1.97	.34	1.76			8.2	1.6	

Figure 4. Measure logit and standard deviation at the output of the item measure for each item on the problem-solving ability test

Based on Figure 4, it is known that the standard deviation value is 0.28. According to the logit measure and standard deviation, the problem solving ability test items can be classified as in the following table:

Table 8. Analysis results of the difficulty level of problem solving ability test items

Measure Logit	Interpretations of The Difficulty Level	Item	Number Of Items
M < -0.28	Very easy	N1(a). N1(e). N2(e). N5(b) dan N6(e)	5
-0.28 ≤ M ≤ 0	Easy	N1(b). N1(c). N1(d). N2(b). N2(d). N3(d). N4(a). N5(a). N5(c). dan N6(b)	10
0 ≤ M ≤ 0.28	Hard	N2(c). N3(a). N3(c). N3(e). N4(c). N4(d). N6(a) dan N6(d)	8
M > 0.28	Very Difficult	N3(b). N4(b). N4(e). N5(d). N5(e). dan N6(c)	6

Based on Figure 4, it is obtained that the level of difficulty of the items in the very easy interpretation is 20% and in the easy interpretation is 33.33%. While the level of difficulty of the items on the difficult interpretation is 26.67% and the very difficult interpretation is 20%. These results indicate that the items on the problem solving ability test instrument are almost evenly distributed at each level of difficulty.

▪ CONCLUSION

The results showed that the problem-solving ability test instrument on the material developed for optical devices had good quality in terms of validity and reliability. Of the six item items, all of them can be used. The difficulty level of the problem solving ability test instrument varies. The problem-solving ability test instrument is appropriate

for use as a measure of problem-solving ability in high school students based on validity, reliability, and level of difficulty.

This research is expected to be a teacher's reference in carrying out problem-solving ability tests on the material of optical instruments and references in compiling problem-solving ability test items. The sample in this study is still limited, it is hoped that future researchers can use a larger sample.

▪ REFERENCES

Apriyani, R., Ramalis, T. R., & Suwarma, I.R. (2019). Analyzing student's problem solving abilities of direct current electricity in stem-based learning. *Journal of Science Learning*. 2(3).85-91

Aristiawan, A., & Istiyono, E. (2020). Developing instrument of essay test to measure the problem-solving skill in physics. *Jurnal Pendidikan Fisika Indonesia*, *16*(2), 72-82.

Bahtiyar, A., & Can, B. (2016). An investigation of problem-solving skills of preservice science teachers. *Educational Research and Reviews Academic Journal*, *11*(23), 2108–2115.

Baig, M., Ali, S. K., Ali, S., & Huda, N. (2014). Evaluation of multiple choice and short essay question items in basic medical sciences. *Pakistan journal of medical sciences*, *30*(1), 3.

Boone, W. J., Yale, M. S., & Staver, J. R. (2014). Rasch analysis in the human sciences.In *Rasch Analysis in the Human Sciences*.

Brookhart, S. M., Walsh, J. M., & Zientarski, W. A. (2006). The dynamics of motivation and effort for classroom assessments in middle school science and social studies. *Applied Measurement in Education*, *19*(2), 151-184.

Chang, C. Y. (2010). Does problem solving= prior knowledge+ reasoning skills in earth science? An exploratory study. *Research in Science Education*, *40*, 103-116.

Darmana, A., Sutiani, A., Nasution, H. A., Sylvia, N. S. A., Aminah, N., & Utami, T. (2021). Analysis of multi rater with facets on instruments hots of solution chemistry based on tawheed. *Journal of Physics: Conference Series*, *1819*(1). https://doi.org/10.1088/1742-6596/1819/1/012038

Dewi, F. H., Samsudin, A., & Chandra, D. T. (2021, November). Developing FDMT to investigate students' mental model on fluid dynamic concept: a Rasch model analysis. In *Journal of Physics: Conference Series* (Vol. 2098, No. 1, p. 012020). IOP Publishing.

Dixon, R. A., & Brown, R. A. (2012). Transfer of learning: connecting concepts during problem solving. *Journal of Technology Education*, *24*(1), 2-17.

Docktor, J. L., Dornfeld, J., Frodermann, E., Heller, K., Hsu, L., Jackson, K. A., & Yang, J. (2016). Assessing student written problem solutions: a problem-solving rubric with application to introductory physics. *Physical Review Physics Education Research*, 12(1).

Docktor, J. L., Strand, N. E., Mestre, J. P., & Ross, B. H. (2015). Conceptual problem solving in high school physics. *Physical Review Special Topics-Physics Education Research*, *11*(2), 020106.

Gok, T., & Silay, I. (2008). The effects of problem-solving strategies on students' achievement, on the cooperative learning groups in physics teaching. *Hacettepe*

Universitesi Egitim Fakultesi Dergisi-Hacettepe University Journal of Education, (34), 116-126.

Güler, N. (2014). Analysis of open-ended statistics questions with many facet rasch model. *Eurasian Journal of Educational Research*, *55*, 73-90.

Haladyna, T. M., & Rodriguez, M. C. (2013). Developing and validating test items.

Heller, K., & Heller, P. (2010). Cooperative problem solving in physics a user's manual. In Tersedia: *http://www.aapt.org/Conferences/newfaculty/upload/Coop-Problem SolvingGuide.* pdf.

Henderson, C., Heller, K., Heller, P., Kuo, V. H., & Yerushalmi, E. (2001). Instructors' ideas about problem solving--setting goals.

Hidayati, F., Tirtawirya, D., Yudhistira, D., & Naviri, S. (2022). The conditioning training program to improve the strength and endurance of football extracurricular participants: content validity and reliability. *Asian Exercise and Sport Science Journal*, *6*(1), 58-66.

Ince, E. (2018). An overview of problem solving studies in physics education. *Journal of Education and Learning*, *7*(4), 191-200.

Istiyono, E., Mustakim, S. S., Widihastuti, W., Suranto, S., & Mukti, T. S. (2019). Measurement of physics problem-solving skills in female and male students by physteproSS. *Jurnal Pendidikan IPA Indonesia*, *8*(2), 170-176.

Kastner, M., & Stangla, B. (2011). Multiple choice and constructed response tests: Do test format and scoring matter?. *Procedia-Social and Behavioral Sciences*, *12*, 263-273.

Kemendikbud. (2016). *Peraturan menteri pendidikan dan kebudayaan republik indonesia nomor 21 tahun 2016* [regulation of the minister of education and culture of the republic of indonesia number 21 of 2016]. Jakarta : Kementrian Pendidikan dan Kebudayaan.

Koçak, D. (2020). Investigation of rater tendencies and reliability in different assessment methods with many facet rasch model. *International Electronic Journal of Elementary Education*, *12*(4), 349–358.

Kubiszyn, T., & Borich, G. D. (2016). Educational testing and measurement. John Wiley & Sons.

Lim, C., & Han, H. (2020). Development of instructional design strategies for integrating an online support system for creative problem solving into a university course. *Asia Pacific Education Review*, *21*(4), 539-552.

Mardapi, D. (2012). Measurement of educational assessment and evaluation. *Yogyakarta: Nuha Medika*, *45*.

Mofreh, S. A. M., Ghafar, M. N. A., Hj Omar, A. H., Mosaku, M., & Ma'ruf, A. (2014). Psychometric properties on lecturers' beliefs on teaching function: Rasch model analysis. *International Education Studies*, *7*(11), 47–55. https://doi.org/10.5539/ies.v7n11p47

N. Nurdini, A. Suhandi, T. Ramalis, A. Samsudin, N. J. Fratiwi, and B. Costu, "Developing multitier instrument of fluids concepts (mifo) to measure student's conception: a rasch analysis approach," *J. Adv. Res. Dyn. Control Syst.*, vol. 12, no. 6, pp. 3069–3083, 2020, doi: 10.5373/JARDCS/V12I6/S20201273.

Ng, S. E., Yeo, K. J., & Mohd Kosnin, A. B. (2018). Item analysis for the adapted motivation scale using rasch model. *International Journal of Evaluation and*

Research in Education (IJERE), *7*(4), 264. https://doi.org/10.11591/ijere.v7i4.15376

OECD. (2003). The PISA 2003 assessment framework – mathematics, reading, science and problem solving, knowledge and skills. OECD.

Permatasari, A. K., Istiyono, E., & Kuswanto, H. (2019). Developing assessment instrument to measure physics problem solving skills for mirror topic. *International Journal of Educational Research Review*, *4*(3), 358-366.

Purwaningsih, E., Sari, S. P., Sari, A. M., & Suryadi, A. (2020). The effect of *stem project based lesrning* and discovery learning on improving students' problem-solving skills of impulse and momentum topic. *Jurnal Pendidikan IPA Indonesia*, 9(4), 465-476.

Rokhmah, A., Sunarno, W., & Masykuri, M. (2017). Science literacy indicators in optical instruments of highschool physics textbooks chapter. *Jurnal Pendidikan Fisika Indonesia*, *13*(1), 19-24.

Scherer, R., & Beckmann, J. F. (2014). The acquisition of problem solving competence: evidence from 41 countries that math and science education matters. *Large-scale Assessments in Education*, *2*(1), 1-22

Sirait, J., Sutrisno, L., Balta, N., & Mason, A. (2017). The development of questionnaire to investigate students attitudes and approaches in physics problem solving. *Jurnal Pendidikan Fisika Indonesia*, *13*(2), 79-87.

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan Rasch pada Assessment pendidikan.* Bandung: Trim Komunikata

Sundari, R., Suhandi, A., & Samsudin, A. (2023). Direct current electric circuit-visualization skills instrument test (dcecvisit) to measure high school student's visualization skills. Jurnal Pendidikan MIPA. *Jurnal Pendidikan*, *24*(1), 84-97.

Suprapto, E., Sumiharsono, R., & Ramadhan, S. (2020). The analysis of instrument quality to measure the students' higher order thinking skill in physics learning. *Journal of Turkish Science Education*, *17*(4), 520-527.

Talib, R., Iahad, N. A., Ashari, Z. M., Rameli, M. R. M., Bakar, Z. A., & Dollah, R. (2019). Rasch strategies for evaluating quality of the Conceptions and Alternative

Tang, T., Vezzani, V., dan Erikson, V. (2020). Developing critical thinking, collective creativity skills and problem solving through playful design jams. *Thinking Skills and Creativity*, 37.

Tumanggor, A. M. R., Jumadi, J., Wilujeng, I., & Ringo, E. S. (2019). The profile of students' physics problem solving ability in optical instruments. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, *5*(1), 29-40.

Uzun, S., Alev, N., & Karal, I. S. (2013). A cross-age study of an understanding of light and sight concepts in physics. *Science Education International*, *24*(2), 129-149.

Waller, E., & Kaye, M. H. (2012). Teaching problem-solving skills to nuclear engineering students. *European Journal of Engineering Education*, *37*(4), 331-342.

Williams, M. (2018). The missing curriculum in physics problem-solving education. *Science & Education*, 27(3), 299-319.

Wilson, F. R., Pan, W., & Schumsky, D. A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. *Measurement and evaluation in counseling and development*, *45*(3), 197-210

Winarto, W., Cahyono, E., & Sumarni, W. (2022). Developing a problem-solving essay test instrument (pseti) in the instruction of basic science concepts in ethnoscience context. *Journal of Turkish Science Education*, *19*(1), 37-51.

Xie, J., & Masingila, J. O. (2017). Examining interactions between problem posing and problem solving with prospective primary teachers: a case of using fractions. *Educational Studies in Mathematics*, *96*(1), 101-118

Yudhistira, D., & Tomoliyus. (2020). Content validity of agility test in karate kumite category. *International Journal of Human Movement and Sports Sciences*, *8*(5), 211–216. https://doi.org/10.13189/saj.2020.080508