
Problem-Solving Skills on Night Sky Brightness and Light Pollution: A Rasch Analysis

Silvia Lutfiani, Winny Liliawati, Ahmad Aminudin

This article has been presented at The 2nd International Physics Conference (IPC)

Universitas Pendidikan Indonesia, Bandung, Indonesia.

August, 16, 2025

Abstract

Light pollution was a global environmental problem that disrupted ecosystems, particularly in the field of astronomy. One of the main indicators of light pollution was the increased brightness of the night sky. This phenomenon was highly relevant to physics education, particularly in Earth and Space Physics courses. However, studies measuring problem-solving skills in the context of real-world scientific phenomena, such as night sky brightness as an indicator of light pollution, were still very limited. Therefore, this study aimed to test the feasibility of a problem-solving skills assessment tool and obtain an overview of students' problem-solving skills in solving problems related to light pollution issues, using night sky brightness as an indicator. This study used a quantitative descriptive design and involved 75 physics education students. The instrument consisted of one main question containing four stages of problem-solving skills. Thus, there were four items analyzed using the Rasch Model. The analysis results showed that the instrument met the assumption of unidimensionality with a very good category. Reliability, reviewed from three aspects: person reliability, item reliability, and Cronbach's Alpha, was categorized as weak, excellent, and sufficient, respectively. All items were deemed appropriate (fit), with the difficulty level consisting of one easy item, two medium items, and one difficult item. All items had excellent discriminatory power. The use of this instrument indicated that, overall, students' problem-solving skills were excellent. This finding indicated that the instrument was suitable for obtaining an overview of students' problem-solving skills. This study implied that the developed instrument could be used in learning to provide an overview of students' problem-solving skills, particularly in the context of light pollution issues.

Keywords: Rasch · Physics · Problem-solving · Night sky brightness · Light pollution

INTRODUCTION

The night sky is a shared and universal heritage for all people. It is an essential part of the environment, particularly in astronomy. Under natural conditions, a dark night sky allows optimal observation of celestial objects without artificial light interference. Light pollution is defined as altering natural light patterns in the night environment caused by introducing

✉ Silvia Lutfiani Winny Liliawati
Silvialutfiani27@upi.edu winny@upi.edu

Ahmad Aminudin
aaminudin@upi.edu

Physics Education Studi Program, Universitas Pendidikan Indonesia, Bandung, Indonesia.

How to Cite: Lutfiani, S., Liliawati, W. & Aminudin, A. (2025). Problem-Solving Skills on Night Sky Brightness and Light Pollution: A Rasch Analysis. *Proceedings of International Physics Conference*, 1(1), 86-101. <https://proceedings.fisikaupi.id/index.php/ipc>

artificial light (Green et al., 2022; Hussein et al., 2021). Approximately 80% of the world's population lives in areas affected by light pollution (Karpińska & Kunz, 2022). It is estimated that 23% of the world's land area is affected by light pollution, and the area illuminated by artificial light increased by 2% annually between 2012 and 2016. (Boyes et al., 2021). In addition, Hussein et al. (2021) also revealed that the use of artificial light is spreading by 6% annually. According to Zheng et al. (2021), several developing countries show a much higher growth rate in the intensity and area of artificial light at night, reaching 10% annually. This increase has direct implications for the deterioration of the condition of the night sky in various regions, one of which is through the phenomenon of light pollution. One of the primary and most widespread forms of light pollution is skyglow. Skyglow is a phenomenon of light glow in the night sky caused by artificial light from the Earth's surface, such as city lights, which is scattered by molecules and aerosols in the atmosphere and then reflected to Earth (Rakhmadi et al., 2023).

To determine the pollution level in an area, night sky brightness can be used as an indicator (Asmoro et al., 2022). The higher the measured sky brightness value, the greater the level of light pollution. This measurement can be carried out using a Sky Quality Meter (SQM) instrument that provides magnitude per square arc second data, or satellite data such as the Visible Infrared Imaging Radiometer Suite (VIRRS) that maps the intensity of artificial light from the Earth's surface. Night sky brightness observations have been carried out in Indonesia using VIRRS and SQM data at 7 LAPAN stations spread across Indonesia (Admiranto et al., 2022). Garut obtained a value of 20.6 mag/arc sec² and a Bortle Scale of 2, which means that light pollution is very low, making it ideal for night sky observations. Next are Biak (20.0 mag/arc sec²), Sumedang (19.6 mag/arc sec²), and Agam (19.5 mag/arc sec²), which are all included in the Bortle Scale 5, indicating fairly dark sky conditions despite a little artificial light interference. In addition, Pasuruan (18.0 mag/arc sec²), including the Bortle Scale 6, indicates the sky is starting to brighten due to human activity. The two locations with the highest levels of light pollution are Pontianak (17.7 mag/arc sec²) and Bandung (17.1 mag/arc sec²), included in the Bortle scale 8, indicating a very bright night sky. In these conditions, star visibility is very low, disrupting astronomical observation activities due to the dominance of artificial light spreading into the atmosphere (skyglow), and dark sky conditions are increasingly difficult to find. Thus, this research focuses on the brightness of the night sky as an indicator of light pollution to provide an overview of the condition of the night sky in a region and to form the basis for efforts to address these challenges.

To face the challenges of light pollution that reduces the natural darkness of the night sky, 21st-century education has an important role to equip the younger generation, especially students, with the competencies needed to understand and analyze these problems, namely, problem-solving skills. Problem-solving skills are a basic activity for humans because humans will continue to face problems in everyday life, so they need the ability to solve them (Hadi & Radiyatul, 2014). There are four stages in problem-solving: understanding the problem, devising a plan, carrying out the plan, and looking back (Polya, 1973; Pradana, 2024). Understanding the problem stage is the ability to explain the problem and identify the main elements of the problem, such as what is being asked and the available data. The devising a plan stage is the ability to design clear steps to solve the problem. The carrying out the plan stage is the ability to implement the plan that has been prepared by checking each step. The last

stage is looking back, which is the ability to re-check the results to ensure that the solution found is correct.

Many studies have shown that problem-solving skills remain a challenge at various levels of education, particularly for Physics Education students (Gök & Silay, 2008; Herayanti et al., 2020; Mashami & Gunawan, 2018; Setyarsih et al., 2018). However, most previous research places these skills within the context of common classical physics phenomena, such as mechanics or thermodynamics. Studies that specifically measure these skills within the context of real-world scientific phenomena, specifically astronomical issues, are still minimal. However, light pollution, with night sky brightness as one indicator, is an astronomical phenomenon that is both relevant and close to everyday life, and suitable for study in Earth and Space Physics courses. By addressing this issue, this study seeks to examine problem-solving skills in physics education by presenting an astronomical context and a global issue. More specifically, this study aims to analyze a problem-solving skills test instrument related to night sky brightness as an indicator of light pollution and describe the level of problem-solving skills of physics education students based on the results of measurements with this instrument. The test is administered after the lesson, so the results are expected to demonstrate students' skills in applying the concepts they have learned. Theoretically, this research can enrich the development of problem-solving skills test instruments based on real-world scientific contexts. Practically, this research can provide an overview of students' problem-solving skills in addressing environmental issues, particularly the issue of light pollution as an indicator of night sky brightness.

METHOD

Research Design

This study employed a quantitative descriptive design, with data analysis encompassing instrument quality testing and descriptive analysis of results. The aim of this study was to analyze the quality of the problem-solving skills test instrument and to obtain an overview of problem-solving skills. The analysis was conducted quantitatively based on the test results given after the completion of the lesson on night sky brightness as an indicator of light pollution in the Earth and Space Physics course.

Population and participants

The population in this study was 75 students studying Earth and Space Physics, a Physics Education study program. The researcher used the entire population as research participants, so the sampling technique used was total sampling. Participants were aged 17-20 years. Of the participants, 54 were female and 21 were male. The study was conducted in the even semester of the 2024/2025 academic year at a university in Bandung City. Bandung City is located in West Java Province, Indonesia. On the map, Bandung City can be seen in Figure 1.



Figure 1. Map of Bandung City (Source from Google Maps)

Research Instrument

This research instrument used a problem-solving skills test in the form of a description of Night Sky Brightness as an Indicator of Light Pollution. The test was administered after students attended the Earth and Space course. The problem-solving skills test consisted of one question covering four problem-solving stages: understanding the problem, devising a plan, carrying out the plan, and looking back. (Ilma et al., 2024; Polya, 1973) . Each problem-solving stage is equipped with an assessment rubric with a scale of 0-3. The details of the questions from the problem-solving skills test are shown in Table 1.

Table 1. Details of Problem-solving Skills Test Question Indicators

No.	Problem-solving Skills Stage	Question Indicator
a	Understanding the problem: Describe the problem and identify the main elements of the problem, such as what is being asked and the data available.	Explain issues related to night sky quality and light pollution based on MPSAS graphs and identify important elements of the graphs presented.
b	Devising a plan: Design clear steps to resolve the problem	Designing graphic analysis steps and atmospheric conditions to determine night sky quality and light pollution
c	Carrying out the plan: Carry out plans based on the steps that have been prepared	Carry out graphic analysis according to plan, determine MPSAS values, classify them with the Bortle Scale, and conclude the quality of the night sky and the level of light pollution
d	Looking back: Double-check the results to ensure that the solution found is correct.	Re-examining the analysis results regarding night sky quality and light pollution levels

Data Analysis Techniques

Instrument Quality

The data obtained from the problem-solving skills test instrument was then analyzed using the Rasch Model. The Rasch Model analysis using WINSTEP software version 5.7.4.0 is reviewed based on the results of item dimensionality, item fit, wright maps, item measure, reliability based on Cronbach's alpha, item reliability, and person reliability values, as well as discriminatory power.

Construct validity in Rasch analysis is tested through dimensionality. Dimensionality analysis uses Table 23 output, namely item dimensionality, which is then interpreted based on Table 2 (Cahyani et al., 2024; Samsudin et al., 2021).

Table 2. Interpretation of Dimensionality

Indicator	Index Value	Interpretation
Raw variance explained by measures	> 0,60%	Very good
	40% – 60%	Good
	20% – 40%	Enough

Reliability in this study was conducted by analyzing item reliability, person reliability, and Cronbach's Alpha. The results of the analysis were obtained from the output of Table 31: Summary statistics, then interpreted based on Table 3 (Fadlan et al., 2024; Sumintono & Widhiarso, 2014).

Table 3. Interpretation of Item Reliability

Indicator	Index Value	Interpretation
Item and Person Reliability	> 0,94	Excellent
	0,91 – 0,94	Very good
	0,81 – 0,90	Good
	0,67 – 0,80	Enough
	< 0,67	Weak
Cronbach's Alpha	> 0,8	Very good
	0,7 – 0,8	Good
	0,6 – 0,7	Sufficient
	0,5 – 0,6	Bad
	< 0,5	Poor

In the Rasch model analysis, item fit refers to three criteria, namely MNSQ outfit, ZSTD outfit, and Pt Measure Corr shown in Table 4 (Aprilia et al., 2020; Boone, 2016; Sumintono & Widhiarso, 2015). If these three criteria are not met by an item, it can be ascertained that the item does not meet the expectations of the Rasch Model and is considered inadequate, so it needs to be improved or replaced (Aprilia et al., 2020; Darma & Sangka, 2025; Rizki et al., 2025).

Table 4. Criteria for Item Fit of Question Items

Indicator	Criteria
Outfit Mean Square (MNSQ)	$0,5 < \text{MNSQ} < 1,5$
Z-standard Outfit (ZSTD)	$-2,0 < \text{ZSTD} < 2,0$
Point-Measure Correlation (Pt. Mean Corr)	$0,4 < \text{Pt. Measure Corr} < 0,85$

The level of difficulty can be determined using the output menu Table 1: Variable (Wright) maps (Purwana et al., 2020). The output analysis results, Table 13: Item measure, are then interpreted by referring to Table 5 (Fadlan et al., 2024; Sumintono & Widhiarso, 2014).

Table 5. Criteria for the Level of Difficulty of Question Items

Criteria	Interpretation
$ME < -1SD$	Easy
$-1SD \leq ME \leq +1SD$	Currently
$ME > +1SD$	Difficult

Discriminatory power is the ability of questions to differentiate students who are able to answer or students who have a high level of ability with students who have low ability in answering questions (Jumini et al., 2023) . The results of the output analysis Table 10: Item Fit are then interpreted based on the results criteria table shown in Table 6 (Fadlan et al., 2024; Sumintono & Widhiarso, 2014) .

Table 6. Criteria for Discriminatory Power of Question Items

Indicator	Index Value	Criteria
Pt. Measure Corr.	$ID \geq 0,40$	Very good
	$0,40 < ID \leq 0,30$	Good.
	$0,30 < ID \leq 0,20$	Not good
	$ID < 0,20$	Bad

Student Problem-solving Skills

This study also used variable analysis (Wright maps) to assess students' problem-solving abilities. This analysis resulted in a map showing the students' high and low abilities and the difficulty levels of different items (Sumintono & Widhiarso, 2015). Furthermore, the data obtained from the problem-solving skills test instrument was further analyzed by calculating scores at each problem-solving stage. The goal was to obtain more specific and detailed student achievement at each stage of the problem-solving process. The method for calculating the percentage of scores obtained per problem-solving stage is as follows.

$$p = \frac{x_i}{x_{max}} \times 100\%$$

Furthermore, the percentage results obtained were interpreted by referring to Table 7 (Choeriyah et al., 2021).

Table 7. Interpretation of Problem-Solving Skills

Percentage (%)	Qualification
$80 \leq X \leq 100$	Very Good
$65 \leq X \leq 79.9$	Good
$55 \leq X \leq 64.9$	Enough
$40 \leq X \leq 54.9$	Not enough

RESULTS AND DISCUSSION

Results

Instrument Quality

The output contains the results of the dimensionality test of the problem-solving skills test instrument. The dimensionality items in WINSTEP software version 5.7.4.0 are shown in Figure 2.

	Eigenvalue	Observed	Expected
Total raw variance in observations =	209.8623	100.0%	100.0%
Raw variance explained by measures =	141.8624	67.6%	62.4%
Raw variance explained by persons =	79.9845	38.1%	35.2%
Raw Variance explained by items =	61.8778	29.5%	27.2%
Raw unexplained variance (total) =	68.0000	32.4%	100.0%
Unexplned variance in 1st contrast =	29.1989	13.9%	42.9%
Unexplned variance in 2nd contrast =	23.0046	11.0%	33.8%
Unexplned variance in 3rd contrast =	15.7966	7.5%	23.2%
Unexplned variance in 4th contrast =	.0002	.0%	.0%

Figure 2. Validity of the Problem-solving Skills Test



Figure 2 shows the results of the item dimensionality test based on the raw variance explained by measures criteria. The results show that the raw variance explained by the measures of the problem-solving skills test with four items is 67.6%.

Figure 3 shows the results of the reliability test of the problem-solving skills test instrument obtained from the output Summary statistics on WINSTEP software version 5.7.4.0.

SUMMARY OF 75 MEASURED (EXTREME AND NON-EXTREME) Person									
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	
MEAN	9.8	4.0	7.30	2.99					
SEM	.2	.0	.49	.42					
P.SD	1.3	.0	4.18	3.57					
S.SD	1.3	.0	4.20	3.60					
MAX.	12.0	4.0	12.53	11.60					
MIN.	5.0	4.0	-10.64	1.43					
REAL RMSE	4.72	TRUE SD	.00	SEPARATION	.00		Person RELIABILITY	.00	
MODEL RMSE	4.66	TRUE SD	.00	SEPARATION	.00		Person RELIABILITY	.00	
S.E. OF Person MEAN	= .49								
Person RAW SCORE-TO-MEASURE CORRELATION = .95									
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .65 SEM = .77									
STANDARDIZED (50 ITEM) RELIABILITY = .00									
SUMMARY OF 4 MEASURED (NON-EXTREME) Item									
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	
MEAN	183.3	75.0	.00	.39	.91	-.62	1.84	.19	
SEM	9.6	.0	1.23	.02	.18	.97	1.28	.86	
P.SD	16.6	.0	2.13	.04	.31	1.69	2.21	1.49	
S.SD	19.2	.0	2.46	.04	.35	1.95	2.56	1.72	
MAX.	204.0	75.0	2.82	.44	1.38	1.84	5.66	2.57	
MIN.	162.0	75.0	-2.66	.35	.57	-2.56	.35	-1.55	
REAL RMSE	.41	TRUE SD	2.09	SEPARATION	5.13	Item	RELIABILITY	.96	
MODEL RMSE	.39	TRUE SD	2.10	SEPARATION	5.39	Item	RELIABILITY	.97	
S.E. OF Item MEAN	= 1.23								

Figure 3. Reliability of the Problem-solving Skills Test

Figure 3 shows the reliability results of the problem-solving skills test instrument reviewed based on person reliability, Cronbach's Alpha value, and item reliability. Person reliability obtained a value of 0.00, which is categorized as weak. Cronbach's Alpha value of 0.65 is categorized as low, and item reliability showed very high values, namely 0.96 and 0.97, which are categorized as excellent.

The output contains the results of the item fit test for the problem-solving skills test. The item fit order in WINSTEP software version 5.7.4.0 is shown in Figure 4.

Item STATISTICS: MISFIT ORDER													
ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR CORR.	AL EXP.	EXACT OBS%	MATCH EXP%	Item
1	204	75	-2.69	.40	1.38	1.85	6.04	2.66	A .63	.75	86.8	86.6	a
4	162	75	2.98	.47	.99	.06	.58	.02	B .57	.57	91.2	91.1	d
2	194	75	-1.35	.35	.72	-1.70	.85	-1.11	b .77	.72	91.2	83.4	b
3	174	75	1.07	.36	.50	-2.93	.31	-1.83	a .71	.65	95.6	85.6	c
MEAN	183.5	75.0	.00	.40	.90	-.68	1.94	.18			91.2	86.7	
P.SD	16.5	.0	2.18	.05	.33	1.81	2.37	1.60			3.1	2.8	

Figure 4. Item Fit of Problem-solving Skills Test

Referring to Figure 4, a summary of the results of the item fit test of the problem-solving skills test is shown in Table 8.

Table 8. Summary of Problem-solving Skills Test Fit Items

Question Items	MNSQ	ZSTD	Pt. Measur Corr	Information
a	6.04	2.66	0.63	In accordance
b	0.85	-0.11	0.77	In accordance
c	0.31	-1.83	0.71	In accordance
d	0.58	0.02	0.57	In accordance

Table 8 shows the results of the item fit test based on three criteria in the Rasch model, namely outfit MNSQ, outfit ZSTD, and Pt. Item a has an MNSQ value of 6.04 and a ZSTD of 2.66, which are outside the criteria, but still has a Pt. Measure Corr value of 0.63, which means it meets the criteria and is still considered appropriate. Item c has an MNSQ value of 0.31, which is outside the criteria, while the ZSTD and Pt. Measure Corr values meet the criteria. Items b and d meet all three criteria with MNSQ values in the range of 0.5-1.5, ZSTD between -2 to +2, and Pt. Measure Corr between 0.4-0.85.

The results of the difficulty level test are obtained from the output variable (Wright) maps in WINSTEP software version 5.7.4.0. The output results are shown in Figures 5 and 6.

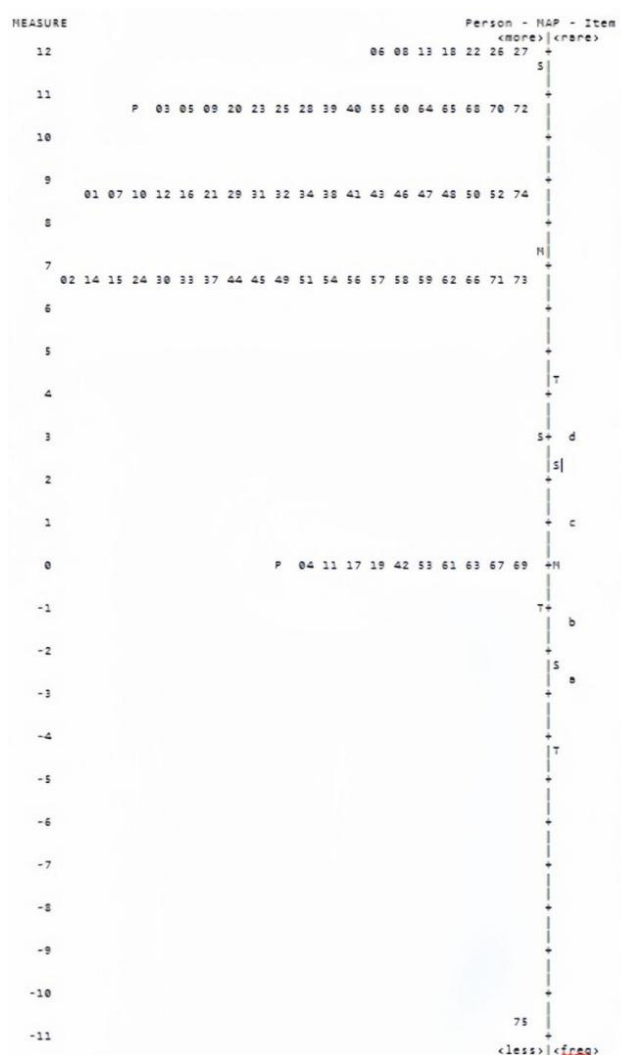


Figure 5. Wright Maps Problem-solving Skills Test

Figure 5 shows the left side of the figure depicting the participants' abilities, while the right side depicts the distribution of logit values for the item difficulty level. The higher the logit value, the more difficult the item is, while the lower the logit value, the easier the item is. The problem-solving skills test shows that the items have varying levels of difficulty, with logit values ranging from 0 to 12. Items with high logit values are located at the top of the map, indicating difficult items, while items with low logit values are located at the bottom of the map and are easy. Items a and b are around -1 logit, indicating that they are both low-difficulty (easy)

items. Items c and d are around +1 to +2 logit, indicating that they are relatively difficult compared to the other items.

The results of the difficulty level test are also obtained from the output—item measure and on WINSTEP software version 5.7.4.0. The output results can be seen in Figure 6.

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ ZSTD	OUTFIT MNSQ ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	Item
4	162	75	2.98	.47	.99 .06	.58 .02	.57	.57	91.2	91.1	d
3	174	75	1.07	.36	.50 -2.93	.31 -1.83	.71	.65	95.6	85.6	c
2	194	75	-1.35	.35	.72 -1.70	.85 -.11	.77	.72	91.2	83.4	b
1	204	75	-2.69	.40	1.38 1.85	6.04 2.66	.63	.75	86.8	86.6	a
MEAN	183.5	75.0	.00	.40	.90 -.68	1.94 .18			91.2	86.7	
P.SD	16.5	.0	2.18	.05	.33 1.81	2.37 1.60			3.1	2.8	

Figure 6. Problem-solving Skills Test Measure Items

Referring to Figure 6, a summary of the test results of the problem-solving skills test items is shown in Table 9.

Table 9. Summary of Problem-solving Skills Test Measure Items

Question Items	Logit value	Interpretation
a	$ME < (-2,18)$	Easy
b	$2,18 \geq ME \geq (-2,18)$	Currently
c	$2,18 \geq ME \geq (-2,18)$	Currently
d	$ME > 2,18$	Difficult

Table 9 shows the results obtained from the analysis of the questions' difficulty level, indicating variations in the level of difficulty of each item. Item a is in the easy category with a value of $-2.69 < 2.18$ logit. Items b and c are in the medium category (between -2.18 to 1.28) with values of -1.35 and 1.07 logit. Item d is in the difficult category because the value is $2.98 < 2.18$ logit.

The results of the discriminatory power test of the problem-solving skills test obtained from the output. The item fit order in WINSTEP software version 5.7.4.0 is shown in Figure 7.

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ ZSTD	OUTFIT MNSQ ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	Item
1	204	75	-2.69	.40	1.38 1.85	6.04 2.66	A .63	.75	86.8	86.6	a
4	162	75	2.98	.47	.99 .06	.58 .02	B .57	.57	91.2	91.1	d
2	194	75	-1.35	.35	.72 -1.70	.85 -.11	b .77	.72	91.2	83.4	b
3	174	75	1.07	.36	.50 -2.93	.31 -1.83	a .71	.65	95.6	85.6	c
MEAN	183.5	75.0	.00	.40	.90 -.68	1.94 .18			91.2	86.7	
P.SD	16.5	.0	2.18	.05	.33 1.81	2.37 1.60			3.1	2.8	

Figure 7. Discriminating Power of Problem-solving Skills Test

Referring to Figure 7, a summary of the results of the discriminatory power of the problem-solving skills test is shown in Table 10.

Table 10. Summary of the Discriminating Power of the Problem-Solving Skills Test

Question Items	Value of Pt. Measur Corr	Interpretation
a	0.63	Very good
b	0.77	Very good
c	0.71	Very good
d	0.57	Very good

Table 10 shows that all the items tested have very good discriminating power because all the Pt. Measur Corr values are at the value $\geq 0,40$.

Problem-solving Skills

Students' skills in answering questions are known using the output menu Table 1: Variable (Wright) maps, namely a menu that also displays the difficulty level of questions analyzed with Rasch, which is shown in Figure 5. Furthermore, the results of the recapitulation of problem-solving skills for each indicator are shown in Figure 8.

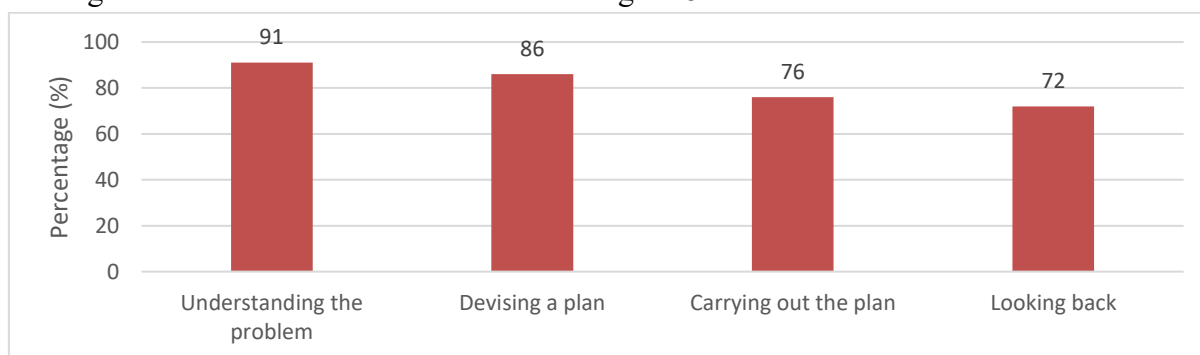


Figure 8. Student Problem-solving Skills

The results of a study of 75 Physics Education Study Program students showed that problem-solving skills in the context of night sky brightness as an indicator of light pollution. Understanding the problem, students obtained a percentage of 91% with a very good category, while at the devising a plan stage, they obtained 86% with a very good category. Meanwhile, at the carrying out the plan stage, the percentage obtained was 76% with a good category, and at the looking back, it was 72% with a good category. Overall, the average problem-solving skill obtained 81% in a very good category.

Discussion

Instrument Quality

Data processing was carried out using WINSTEP software version 5.7.4.0. The quality of the instrument was reviewed from the results of dimensionality (item dimensionality), reliability (person reliability, item reliability, and Cronbach's Alpha), the level of suitability of the items (item fit), the level of difficulty of the items reviewed from Wright maps and item measures, and the discriminating power.

The results of the dimensionality test as a construct validity test showed that the raw variance explained by the measure's value for the problem-solving skills test instrument with four questions was 67.6%. Referring to the requirements proposed by Cahyani et al. (2024) and Samsudin et al. (2021), which stipulate that the minimum raw variance explained by the measure's value is 20%, the problem-solving skills test instrument is declared to meet the

unidimensionality assumption with a very good category. This aligns with the findings of Cahyani et al. (2024), who obtained a score of 66.8% in the very good category, namely, truly measuring the construct. The high value of the raw variance explained by measures obtained indicates that the questions have a high suitability with the construct that is the focus of measurement, in this study, problem-solving skills. With the fulfillment of unidimensionality, the interpretation of test results can accurately represent the level of problem-solving skills of participants.

The reliability results of the problem-solving skills test instrument were reviewed based on person reliability, item reliability, and Cronbach's Alpha value. Person reliability obtained a value of 0.00 in the weak category, item reliability was 0.96 in the excellent category, and a Cronbach's Alpha value of 0.65 in the sufficient category. This aligns with the research of Samsudin et al. (2021), who obtained the results of person reliability obtained a value of 0.58 in the weak category, item reliability, 0.89 in the good category, and Cronbach's Alpha, 0.60 in the sufficient category. In their research, Samsudin et al. (2021) stated that the instrument they developed was reliable. Then, in line with the research of Purwana et al. (2020) who analyzed three items in their instrument, obtained the results of person reliability 0.28 in the weak category, item reliability 0.00 in the weak category, and a Cronbach's Alpha value of 0.58 in the poor category indicating a lack of interaction between the person and the item as a whole. This condition may be influenced by the relatively small number of questions and participants. The greater the number of questions and participants, the higher the reliability, reaching the very good category. In this study, due to the large number of participants (75), the low personal reliability was due to the limited number of questions. The problem-solving skills test instrument was deemed reliable when referring to the same decision-making criteria as in Samsudin et al.'s (2021) study.

The item fit test results are based on three criteria in the Rasch model: outfit MNSQ, outfit ZSTD, and Pt. Item a has an MNSQ value of 6.04 and a ZSTD of 2.66, which are outside the criteria, but still has a Pt. Measure Corr value of 0.63, which means it meets the criteria and is still declared appropriate. Item c has an MNSQ value of 0.31, outside the criteria, while the ZSTD and Pt. Measure Corr values meet the criteria. Items b and d meet all three criteria with MNSQ values of 0.5 to 1.5, ZSTD between -2 to +2, and Pt. Measure Corr between 0.4-0.85. These conditions indicate that the item can detect extreme deviant responses, can distinguish between high and low participant abilities, and there is no need for any items to be repaired or replaced. In line with the research of Aqida et al. (2023), all items in the instrument they developed are suitable for the Rasch model. This is proven by all fit criteria, including MNSQ outfit value of 0.5 to 1.5, ZSTD outfit value of -2 to +2, and positive Pt. Measure the Corr value, and declare it to meet the item fit criteria. Thus, all items of the problem-solving skills test instrument are by the expectations of the Rasch model.

The results of the analysis of the level of difficulty through the item measure show that item A is in the easy category with a value of $-2.69 < 2.18$ logit. Items b and c are in the medium category (between -2.18 to 1.28) with values of -1.35 and 1.07 logit. Item d is in the difficult category because the value is $2.98 < 2.18$ logit. Therefore, the problem-solving skills test instrument consists of one easy item, two medium items, and one difficult item. This shows that the level of difficulty of the questions shows variations in the level of difficulty in each item. According to Jumini et al. (2023), a good question is a question that is not too difficult or too

easy. In line with the research of Lumbanraja & Daulay (2018) in developing their test instrument, the distribution of easy questions was 20%, medium questions 60%, and difficult questions 20%. This condition shows that medium questions dominate, and easy and difficult questions are relatively fewer. Thus, all the questions on the problem-solving skills test instrument place most of the questions in the medium category as part of the difficulty level balance.

The results of the discriminatory power test showed that all the items tested had very good discriminatory power because all the Pt. Measur Corr values were at the value $\geq 0,40$. This condition indicates that the questions were able to differentiate participants who were able to answer the questions or participants who had a high level of ability from participants who had low ability to answer the questions (Jumini et al., 2023), in line with Elfina's (2023) research, which conducted a discriminatory power analysis with Pt. Measur Corr results between 0.4 and 0.85, indicating good and acceptable discriminatory power. The same thing with Sulsilah et al. (2023), showing that all the instrument items had Pt. Measur Corr values in the range of 0.4 to 0.85 and met the Rasch model fit criteria.

Problem-Solving Skills

Participants' problem-solving skills were identified using Table 1: Variable (Wright) maps output menu, displaying the participants' abilities analyzed using Rasch. The output results can be seen in Figure 5. The Person-Map (Wright Map) confirms the instrument's ability to differentiate respondents' competency levels (Thresia et al., 2025). The results indicate that most participants were around 17 to 11 logits, exceeding most item difficulty levels. This position is well above the difficulty level of most items, indicating that students have good abilities in solving the problem-solving skills tested. Furthermore, in the study, a small number of participants were at a moderate to low ability level (0 logit and below), but their number was relatively small. This group was at the same level or below some items, indicating they may have difficulty with certain items, especially those in the moderate to high difficulty category. This information is important to illustrate that although most participants mastered problem-solving skills, some individuals still needed further attention or reinforcement of the material. The instrument contained questions with varying levels of difficulty (easy, medium, and difficult). However, the homogeneity of participants' abilities weakened the differences in answers between individuals, resulting in a weak person reliability value. This was because the instrument was administered after the learning process was completed, so most participants had mastered the material on night sky brightness as an indicator of light pollution. Thus, the results showed that most participants were above the majority of items, indicating good mastery of problem-solving skills. This condition is in line with research by Sari & Muhtadi (2025), which showed that participants' skills were better than the level of difficulty of the questions developed.

After obtaining a general overview of the participants' problem-solving skills through the Wright Map, the next step was identifying which indicators had been mastered well and which still needed strengthening. Understanding the problem stage, participants obtained a percentage of 91% with a very good category, while at the devising a plan stage, they obtained 86% with a very good category. Meanwhile, at the carrying out the plan stage, the percentage obtained

was 76% with a good category, and at the carrying out the plan and looking back, it was 72% with a good category.

The understanding of the problem stage obtained the highest percentage. At this stage, most participants could present the problem clearly and mention all important elements such as known and asked data. That is driven by group learning activities that require students to understand the problem deeply, such as determining the direction and purpose of the problem presented, identifying important data through data collection related to the night sky's brightness, and defining or formulating the problem specifically. In line with Unaenah et al. (2023), learning with group activities can improve students' ability to identify important elements and formulate problems well, thereby increasing the percentage of problem understanding.

The next highest percentage was obtained at the devising a plan stage, in this stage, most students could design steps for solving and link them to known data. That is because students are trained to identify and develop various alternatives to address previously encountered problems in their learning. Therefore, this stage continues the previous stage, from understanding the problem to planning a solution strategy. Because some students understood the problem very well in the previous stage, this stage does not appear to be difficult for them. According to Uluçınar (2023), when students understand a problem, they can generate hypotheses about possible solutions.

The carrying out of the plan and looking backstage was the indicator with the lowest percentage compared to the other stages. This means students' difficulties at this stage were higher than at other stages. In learning, carrying out the plan is related to selecting the best idea from various alternatives generated to be processed into a concrete solution to light pollution, including systematic steps in implementing the chosen solution. Viewed from the aspect of working on the problem, at the carrying out the plan stage, most participants were able to solve the problem according to the plan, determine the MPSAS value, and classify it using the Bortle Scale. However, they did not thoroughly analyze the results. In line with Arsyabinta et al. (2023), students' difficulties at the stage of carrying out the plan were caused by a mismatch between the plan and the problem-solving, an unsystematic explanation of the work method, and incomplete writing of the solution steps. According to Mahisna et al. (2022), difficulties in carrying out the plan can be caused by a lack of conceptual understanding and anxiety that hinder the work process.

Furthermore, in the looking-back stage, participants could obtain analytical results related to the quality of the night sky and the level of light pollution. However, they could not fully articulate the reasons supporting the conclusions drawn because the analysis in the previous stage was limited. In the learning process, this stage is related to assessing the supporting or inhibiting factors of the chosen solution, or, in other words, related to reflection and verification of the solution. Because students experienced difficulties in carrying out the plan, this impacted the next activity, namely, looking back, which was also suboptimal. According to Enlisia et al. (2020), difficulties occurred because students did not know how to look back correctly and were reluctant to double-check. In accordance with research by Sari & Muhtadi (2025), it was revealed that the looking-back stage was the stage with the highest difficulty or errors, and looking back had the lowest level of mastery or application. Nevertheless, overall, the average

problem-solving skills obtained were 81%, categorized as very good. This indicates that most participants could complete each problem-solving stage very well.

CONCLUSION

The analysis results show that the instrument meets the unidimensionality assumption with a very good category. Reliability, reviewed from three aspects: person reliability, item reliability, and Cronbach's Alpha, is in the weak, excellent, and sufficient categories, respectively. All items are declared appropriate (fit), with the difficulty level consisting of one easy item, two medium items, and one difficult item. All items have a discriminatory power in a very good category. The use of this instrument, overall, shows that students' problem-solving skills are in a very good category. This finding indicates that the instrument is suitable for use to obtain an overview of students' problem-solving skills. Educators can use this instrument to obtain an overview of problem-solving skills and identify students' weaknesses at certain stages in the problem-solving process, enabling educators to design more appropriate learning strategies. This study has a limited number of items, so future research is recommended to increase the number of items to improve person reliability and Cronbach's Alpha value, thereby supporting the overall quality of the instrument.

REFERENCES

- Admiranto, A. G., Priyatikanto Rhorom, Maryam, S., Elyyani, Kurniawati, S., & Saputro, M. F. E. (2022). Light Pollution Observations in Indonesia. In L. Hufnagel (Ed.), *Light Pollution, Urbanization and Ecology*. IntechOpen. <https://doi.org/10.5772/intechopen.96897>
- Aprilia, N., Susilaningsih, E., Priatmoko, S., & Kasmui. (2020). Desain Instrumen Tes Pemahaman Konsep Berbasis HOT dengan Analisis Model Rasch. *Chemistry in Education*, 9(2), 1–8.
- Aqida, D. S., Destia Khairunisa, Meisya, R., & Jannah, R. (2023). The Use of Rasch Model in Analyzing Grade V Elementary School Mathematics Items. *Sunan Kalijaga International Journal on Islamic Educational Research*, 7(1), 57–67. <https://doi.org/10.14421/skijier.2023.71.06>
- Arsyabinta, W., Surya, A., & Winarni, R. (2023). Analisis Kesulitan Pemecahan Masalah Matematika Berdasarkan Teori Polya pada Peserta Didik Kelas IV Sekolah Dasar. *Didaktika Dwija Indria*, 11(4), 13–18. <https://doi.org/10.20961/ddi.v11i4.76962>
- Asmoro, C. P., Susanti, H., & Utama, J. A. (2022). DIY Photometer for Night Sky Brightness in Earth & Space Laboratory. *Jurnal Riset Dan Kajian Pendidikan Fisika*, 9(2), 68–75. <https://doi.org/10.12928/jrpkpf.v9i2.167>
- Boone, W. J. (2016). Rasch Analysis for Instrument Development: Why, When, and How? *CBE—Life Sciences Education*, 15(4), rm4. <https://doi.org/10.1187/cbe.16-04-0148>
- Cahyani, N. A. P., Suryana, D., & Sunarya, Y. (2024). Pengembangan Instrumen Citra Tubuh: Analisis Rasch Model. *Jurnal Psikologi*, 20(2), 132. <https://doi.org/10.24014/jp.v20i2.21041>
- Choeriyah, Yulianti, D., & Yulianti, I. (2021). Pengembangan Lembar Kerja Siswa (LKS) Fisika Berbasis Problem Based Learning Untuk Meningkatkan Kemampuan Pemecahan Masalah. *UPEJ (Unnes Physics Education Journal)*, 10(1). <https://doi.org/https://doi.org/https://doi.org/10.15294/upej.v10i1.46756>
- Darma, I. K., & Sangka, I. G. N. (2025). Validation of Applied Mathematics Exam Test Using Rasch Model Approach: Case Study in Diploma 3 Study Program of Refrigeration and Air Conditioning Engineering Politeknik Negeri Bali. *International Journal of Current Science Research and Review*, 08(01), 33–44. <https://doi.org/10.47191/ijcsrr/v8-i1-05>
- Elfina, S. (2023). Analisis Butir Soal Penilaian Akhir Semester (PAS) pada Mata Pelajaran PPKKn Menggunakan Aplikasi Winstep dan Item and Tes Analysis (Iteman) 4.3. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, 8(3), 415–430.



- Enlisia, A. P., Rahardjo, S., & Sisworo, S. (2020). Kesulitan Siswa Dalam Menyelesaikan Masalah Matematika Berdasarkan Langkah Polya. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, 5(12), 1820. <https://doi.org/10.17977/jptpp.v5i12.14347>
- Fadlan, M., Fadillah, N., Liliawati, W., & Mustika Sari, I. (2024). Pengembangan Instrumen Sound Wave Two Tier (Sowatt) Diagnostic Test: Analisis Rasch Model. *U-Teach: Journal Education of Young Physics Teacher*, 5, 81–90. <https://jsr.unuha.ac.id/index.php/u-teach>
- Gök, T., & Silay, L. (2008). Effects of Problem-Solving Strategies Teaching on Problem-Solving Attitudes of Cooperative Learning Groups In Physics Education. *Journal of Theory and Practice in Education*, 4(2), 253–266. http://eku.comu.edu.tr/index/4/2/tgok_isilay.pdf
- Herayanti, L., Widodo, W., Susantini, E., & Gunawan, G. (2020). The Effectiveness of Blended Learning Model Based on Inquiry Collaborative Tutorial Toward Students' Problem-Solving Skills in Physics. *Journal for the Education of Gifted Young Scientists*, 8(3), 959–972. <https://doi.org/10.17478/JEGYS.675819>
- Ilma, K. N., Maulidah, R., & Susanti, E. (2024). The Effect of the Process-Oriented Guided Inquiry Learning Model on Students' Problem-Solving Skills in the First Law of Thermodynamics. *Jurnal Penelitian Pendidikan IPA*, 10(11), 8686–8692. <https://doi.org/10.29303/jppipa.v10i11.4819>
- Jumini, S., Madnasri, S., Cahyono, E., & Parmin, P. (2023). Analisis Kualitas Butir Soal Pengukuran Literasi Sains Melalui Teori Tes Klasik Dan Rasch Model. In *Prosiding Seminar Nasional Pascasarjana*.
- Lumbanraja, L. H., & Daulay, S. (2018). Analisis Tingkat Kesukaran dan Daya Pembeda pada Butir Tes Soal Ujian Tengah Semester Bahasa Indonesia Kelas XII SMA Negeri 7 Medan Tahun Pembelajaran 2016/2017. *Kode: Jurnal Bahasa*, 6(1), 15–24.
- Mahisna, A. B., Sary, R. M., & Cahyadi, F. (2022). Kesulitan Peserta Didik Dalam Pemecahan Masalah Matematika Berdasarkan Prosedur Polya Pada Materi Perkalian Dan Pembagian Di Kelas Ii Sd. *Journal on Mathematics Education Research (J-MER)*, 3(2), 99–132. <https://doi.org/10.17509/j-mer.v3i2.51393>
- Mashami, R. A., & Gunawan, G. (2018). The Influence of Sub-Microscopic Media Animation on Students' Critical Thinking Skills Based on Gender. *Journal of Physics: Conference Series*, 1108(1). <https://doi.org/10.1088/1742-6596/1108/1/012106>
- Polya, G. (1973). *How to Solve It*. Princeton University Press.
- Pradana, L. N. (2024). Problem-solving Strategy: Mathematical Problem-solving Model Within the Polya' Framework. *KnE Social Sciences*. <https://doi.org/10.18502/kss.v9i6.15327>
- Purwana, U., Rusdiana, D., & Liliawati, W. (2020). Pengujian Kemampuan Menginterpretasikan Grafik Kinematika Calon Guru Fisika: the Polytomous Rasch Analysis. *ORBITA: Jurnal Kajian, Inovasi Dan Aplikasi Pendidikan Fisika*, 6(2), 259. <https://doi.org/10.31764/orbita.v6i2.3264>
- Rizki, N., Suryadi, S., Nurmegawati, L., Eneng, S., & Mu, S. (2025). Rasch Model for Analysis of Scientific Attitude Instruments in the Context of Secondary School Science Education. *Equator Science Journal (ESJ)*, 3(2), 98–106. <https://doi.org/10.61142/esj.v3i2.234>
- Samsudin, A., Cahyani, P. B., Purwanto, Rusdiana, D., Efendi, R., Aminudin, A. H., & Coştu, B. (2021). Development of a multitier open-ended work and energy instrument (MOWEI) using Rasch analysis to identify students' misconceptions. *Cypriot Journal of Educational Sciences*, 16(1), 16–31. <https://doi.org/10.18844/cjes.v16i1.5504>
- Sari, R., & Muhtadi, D. (2025). Analisis Kesalahan Siswa pada Soal Garis dan Sudut berdasarkan Tahapan Polya. *Jurnal Inovasi Pembelajaran Matematika: PoweMathEdu*, 04(02), 377–388.
- Setyarsih, W., Azninda, H., & Jatmiko, dan B. (2018). Identifikasi Awal Kemampuan Pemecahan Masalah Fisika Mahasiswa Pendidikan UNESA, Kemampuan Metakognisi Fisika, Sikap dan Strategi Pemecahan Masalahnya. *Seminar Nasional Fisika (SNF) 2018*.
- Sulsilah, H., Hidayat, A., & Ramalis, T. R. (2023). Analysis of Computational Thinking Instrument for High School Student Using Rasch Model. *Jurnal Penelitian Pendidikan IPA*, 9(3), 1445–1450. <https://doi.org/10.29303/jppipa.v9i3.2771>

- Sumintono, B., & Widhiarso, W. (2014). *Aplikasi Model Rasch Untuk Penelitian Ilmu-Ilmu Sosial* (B. Trim (ed.); Revisi). Trim Komunikata Publishing House.
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch: pada Assessment Pendidikan* (B. Trim (ed.); Issue September). Trim Komunikata.
- Thresia, F., Raja, P., Nurweni, A., & Kadaryanto, B. (2025). Development and Validation of the Multiculturally Responsive Teaching (MRT) Instrument: A Rasch Model Analysis for Assessing Multicultural Competence in English Education. *Tadris: Jurnal Keguruan Dan Ilmu Tarbiyah*, 10(1), 303–314. <https://doi.org/10.24042/tadris.v10i1.26547>
- Uluçinar, U. (2023). The Effect of Problem-Based Learning in Science Education on Academic Achievement: A Meta-Analytical Study. *Science Education International*, 34(2), 72–85. <https://doi.org/10.33828/sei.v34.i2.1>
- Unaenah, E., Sari, A. P., Aini, F. A., & Widiati, W. (2023). Analisis Faktor-Faktor yang Mempengaruhi Kemampuan dalam Memecahkan Masalah Matematika di Tingkat Sekolah Dasar. *Jurnal Pendidikan : SEROJA*, 2(3).