23 (4), 2022, 1701-1716

Jurnal Pendidikan MIPA

e-ISSN: 2550-1313 | p-ISSN: 2087-9849

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

MoIP-ST: An Instrument to Measure Students' Problem Solving Skills on Momentum and Impulse

Itsna Rona Wahyu Astuti[1,*], Ida Kaniawati[1], Achmad Samsudin[1], Endi Suhendi[1], Supardiyono[2], Andreas Suparman[1], Rona Hegarna[1], & Rosi Vera[1]

[1]Department of Physics Education, Universitas Pendidikan Indonesia, Indonesia

[2]Department of Physics, Universitas Negeri Surabaya, Indonesia

Abstract: This research aims to develop the MoIP-ST (Momentum and Impulse Problem-Solving Test) instrument to measure students' problem-solving abilities according to the stages of problem-solving skills. The research method used is 4). The research participants were 64 students, consisting of 32 rural and 32 urban 10th-grade students in East Java, who were randomly selected. The first analysis is a multi-rater Rasch measurement based on the results of expert validation, which indicates that item I1 needs improvement. The second analysis is the analysis of the MoIP-ST instrument using the Rasch model, which shows that overall the items on the MoIP-ST are valid and reliable with a Cronbach Alpha of 0.87. The third analysis is the use of the MoIP-ST, which shows that the problem-solving ability of rural students is lower than that of urban students. Student problem-solving needs to be improved through models, methods, learning models, and teaching materials that support problem-solving.

Keywords: test instrument, problem solving skills, physics learning.

Abstrak: *Penelitian ini bertujuan untuk mengembangkan instrumen MoIP-ST (Momentum and Impulse Problem-Solving Test) untuk mengukur kemampuan memecahkan masalah siswa sesuai dengan tahapan kemampuan memecahkan masalah. Metode yang digunakan penelitian adalah 4). Partisipan penelitian adalah 64 siswa yang terdiri dari 32 siswa pedesaan dan 32 siswa perkotaan kelas 10 di Jawa Timur yang dipilih secara acak. Analisis pertama adalah pengukuran Rasch multi-rater dari hasil validasi ahli menunjukkan bahwa item I1 perlu perbaikan. Analisis kedua adalah analisis instrumen MoIP-ST menggunakan Rasch Model menunjukkan secara keseluruhan item pada MoIP-ST valid dan reliabel dengan Cronbach Alpha sebesar 0,87. Analisis ketiga adalah penggunaan MoIP-ST menunjukkan bahwa kemampuan pemecahan masalah siswa pedesaan lebih rendah dibandingkan siswa perkotaan. Pemecahan masalah siswa perlu ditingkatkan melalui model, metode, model pembelajaran, dan bahan ajar yang mendukung pemecahan masalah.*

Kata kunci: *instrumen tes, keterampilan pemecahan masalah, pembelajaran fisika.*

▪ INTRODUCTION

The ability to solve problems is one of the skills required to face challenges in the twenty-first century (Dewi et al., 2019). Problem-solving abilities must be emphasized in the student learning process because they are an important component of learning and are regarded as fundamental in physics learning (Dewi et al., 2019; Sutarno et al., 2017). Problem-solving is defined as a set of procedures that must be followed to solve problems, as well as complex cognitive activities that represent mental constructs that must be correctly understood and solved (Dewi et al., 2019; Jonassen, 2011; Polya, 1957). Furthermore, problem-solving abilities are defined as a person's ability to find solutions through processes involving information gathering and organization (Sujarwanto et al., 2014; Sutarno et al., 2017). Problem-solving begins by recognizing

Itsna Rona Wahyu Astuti
*Email: itsnarona@upi.edu

DOI: http://dx.doi.org/10.23960/jpmipa/v23i4.pp1701-1716
Received: 05 November 2022
Accepted: 20 December 2022
Published: 25 December 2022

that a problem exists and developing an appropriate understanding of the problem. The ability to solve problems is required to identify specific problems, plan and implement solutions, and analyze and evaluate the solutions provided (OECD, 2014). Problem-solving is heavily reliant on students' prior experiences and perspectives on the issue at hand (Schoenfeld, 1992). Several factors influence problem-solving abilities, including the structure of knowledge and the nature of the problems encountered (Chi et al., 1981). The context of students' knowledge also influences their ability to solve problems (Docktor et al., 2015). Solving physical problems effectively requires students to identify, determine and solve problems using logic, literature, and creative thinking (Hegde & Meera, 2012). Problem-solving skills help students think about how to solve problems using relevant concepts and theories (Dewi et al., 2019). Students' knowledge, skills, and understanding can be applied to problem-solving (Dewi et al., 2019; Kohl & Finkelstein, 2008).

Following on from the previous explanation, learning physics necessitates strong problem-solving abilities. Several treatments based on learning models, as well as the development of instruments and media, can be used to improve physics problem-solving skills (Masitoh et al., 2021). Problem-solving instruments are focused on a specific problem that needs to be solved. The structure of a problem is referred to as its problem characteristics. Problems are classified as structured or unstructured based on their structure (Sutarno et al., 2017). A well-structured problem contains all of the information required to solve the problem, which necessitates the use of several ordered and perspective approaches, convergent answers, and defined solutions. Whereas problems in everyday life are more irregular and undefined, the elements of the problem are usually unknown, necessitate different solutions, have different criteria for evaluating solutions, and lack the certainty of concepts and principles required for solving and organizing issue (Gok, 2010). However, before students can solve unstructured problems, they must first become accustomed to solving structured problems so that they are accustomed to problem-oriented questions and can solve them correctly.

Most of the test instruments for problem-solving abilities are in the form of multiple-choice tests (J. Sirait et al., 2017; Nadapdap & Istiyono, 2017). However, multiple-choice tests have drawbacks, so students depend on the answer choices (Kubiszyn & Borich, 2013). In addition, multiple-choice problem-solving tests have a great opportunity for students to make guesses, so the test results are not able to measure the problem-solving ability that should be measured. As a result, students are only able to answer the questions correctly, but they do not know the concepts related to the problems in these questions (Henderson et al., 2016) In addition, multiple-choice tests only contain students' final answers without providing complete information regarding students' problem-solving abilities (Kastner & Stangl, 2011). Therefore, we need a test that can provide information related to students' problem-solving abilities and that covers each stage of problem-solving carried out by students.

Open-ended essay tests are very suitable for representing students' high-level cognitive skills (Haladyna & Rodriguez, 2013). Open-ended essay tests are considered to measure students' complex abilities when students are given the freedom to compile, interpret, integrate, and use knowledge and information in solving new problems or being original and innovative in problem-solving (Kubiszyn & Borich, 2013). Open-ended essay tests are used because answering questions requires students to use higher-order thinking skills (Baig et al., 1969). However, if the problem-solving ability test is only open-ended without guiding students through the stages of problem-solving, it will allow students to only answer the final results. Therefore, this study aims to develop an open-ended test instrument that is capable of measuring problem-solving abilities by loading structured problems and paying attention to the stages of problem-solving. This instrument is named MoIP-ST (Momentum and Impulse Problem Solving Test), which can be used to measure students' problem-solving abilities according to the stages of problem-solving abilities related to physics problems.

▪ METHOD

Participants

This research included 64 grade 10 high school students in the even semester 2020/2021 who had received momentum and impulse material in East Java. The students were divided into 32 rural students and 32 urban students. Students were chosen using random sampling.

Research Design and Procedures

The research method used in this research is 4D. The four stages of the 4D model are as follows: (1) defining the MoIP-ST, (2) designing the MoIP-ST, (3) developing the MoIP-ST, and (4) socializing the MoIP-ST. The definition stage consists of conducting literature studies related to problem-solving analyze the curriculum as well as material momentum and impulse. The design phase consists of making item indicators, item questions, answer keys, score guidelines, and design content for each item. The development stage is to construct the MoIP-ST design based on the initial design. The MoIP-ST instrument was then validated by three experts. Each expert evaluates each item. The results of expert validation were then analyzed using multifaceted Rasch measurements. The dissemination stage was used to test the MoIP-ST instrument with students. The results of the implementation of the MoIP-ST instrument were analyzed using the Rasch model. The analysis used is to measure the validity of MoIP-ST, the reliability of MoI-PST, the item difficulty level of MoIP-ST, and lay students' problem-solving skills.

Instruments

The instrument used in this study is an instrument whose stages correspond to the stages of problem-solving. This instrument consists of seven open-ended questions relating to momentum and impulse material. The distribution of item indicators relates to the topic of momentum and impulse as presented in Table 1.

Table 1. Item indicator distribution

Concepts	Item Number
Analyze the relationship between momentum and impulse in everyday life	1.2.4
Apply the law of conservation of momentum in solving the problem of collision events	3.6.7
Analyze the coefficient of restitution in the event of a collision	5

Each of these items is developed by following the four stages of problem-solving, which consist of (1) exploring and understanding the problem, (2) diagnosing the problem, (3) planning and implementing plans, and (4) monitoring and reflecting on solutions. Each stage of problem-solving has indicators that must be fulfilled by students. The indicator was developed from problem-solving stages that could be adapted for solving problem-based questions. Problem-solving indicators are described in the design stage. Thus, the problem-solving instrument related to the topic of momentum and impulse is given the name MoIP-ST (Momentum and Impulse Problem-Solving Test).

Data Analysis

The analysis used in this research consists of four stages of analysis. The first step in the analysis is to analyze the results of expert validation. Expert validation results were analyzed using the multi-rater Rasch measurement. The second analysis is to analyze the items in the MoIP-ST that have been developed. Instrument analysis using the Rasch Model. Instrument analysis was conducted to determine the validity, reliability, and difficulty level of the items. Instrument analysis was carried out using Rasch analysis. The validity of the instrument is reviewed based on the suitability of each item. Item validity is determined from the output of the Item Fit Order on outfit mean square (MNSQ), outfit Z-Standard (ZSTD), and point measure correlation (PT MEASURE CORR) (Sumintono & Widhiarsho, 2015). In addition, instrument validity is also determined from the output of unidimensionality in the value of raw variance explained by measures. The Rasch analysis was also used to assess instrument reliability. Rasch's analysis yields several values, including person reliability, item reliability, and Cronbach Alpha. The consistency of students' answers is demonstrated by person reliability. Item reliability reflects the instrument's item quality. Cronbach's alpha, on the other hand, depicts the interaction between the person and the item as a whole. The third analysis is to analyze the students' problem-solving skills based on output variable (wright) maps. The fourth analysis is to analyze the students' problem-solving at each stage of problem-solving.

▪ RESULT AND DISSCUSSION

Analysis of the results of each stage of 4D model development (defining, designing, developing, and disseminating) on the MoIP-ST will be discussed in detail below.

Defining MoIP-ST

The defining stage in this research is in the form of literature studies related to problem-solving and curriculum analysis as well as momentum and impulse material analysis. The results of the literature study show that problem-solving competence is an individual's capaurban to engage in cognitive processing to understand and solve problem situations. Problem-solving begins with acknowledging that there is a problem situation and building understanding accordingly. Problem-solving skills are needed to identify specific problems to be solved, plan and implement solutions, and monitor and evaluate the solutions provided. The problem-solving ability has several stages consisting of (1) exploring and understanding; (2) representing and formulating; (3) planning and executing; and (4) monitoring and reflecting. Thus, stages of problem-solving in problem-solving requires four stages consisting of (1) exploring and understanding the problem; (2) diagnosing the problem; (3) planning and implementing plans; and (4) monitoring and reflecting on solutions (OECD, 2014; Sujarwanto et al., 2014).

Designing MoIP-ST

Designing the MoIP-ST instrument involves making question indicators, and assessment indicators as score guidelines, item questions, and answer keys. Design an assessment indicator at each stage that is following solving the problem based on the stages of problem-solving. Table 2 shows the indicators for each stage.

Table 2. Stages and indicators of problem-solving

No.	Stages	Indicators for each stage
1.	Exploring and understanding the problem	• Identify the problem according to the concept • Make a list of known quantities • Determine the amount in question
2.	Diagnosing the problem	• Make a diagram, sketch, or symbol that describes the problem • Determine the relationship between quantities
3.	Planning and implementing plans	• Determine the right equation for problem-solving • Substituting the magnitude values into the equation • Perform calculations using the selected equation
4.	Monitoring and reflecting on solutions	• Evaluate the suitability of the solution to the concept • Evaluating units

The stages and indicators shown in table 1 are used as a guide in the following stage. Each item includes the four stages of problem-solving. And each stage of problem-solving performed by students is graded based on these indicators. As a result, each stage of problem-solving on each item has a fixed score.

Developing MoIP-ST

During the development stage, the instrument was prepared according to the initial design of the MoIP-ST instrument. The MoIP-ST consists of seven open-ended questions arranged according to the stages of problem-solving. Each question goes through four stages of problem-solving, which consist of (1) exploring and understanding the problem; (2) diagnosing the problem; (3) planning and implementing plans; and (4) monitoring and reflecting on solutions. Open-ended questions allow students to write independently regarding concepts and steps for solving problems related to problem-solving. Figure 1 illustrates one of the items on the MoIP-ST instrument.

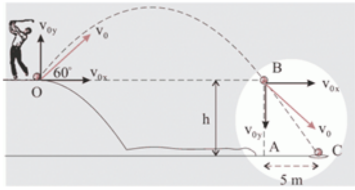
Indicator of Item	Item Question	Stages of Problem-Solving	Scores	Indicator of Each Stage
Analyze the relationship between momentum and impulse in everyday life.	 A golfer wants to insert a ball with a mass of 0.2 kg (g=10 m/s²) which will be inserted into hole C as shown in the figure. It takes 0.01 seconds for the bat to touch the ball. The ball travels the path B-C in 1 s. Determine the magnitude of the force needed by the golfer to hit the ball so that it goes right into hole C!	**Exploring and understanding the problem** Given: Waktu B-C = 1 s X_{AC} = 5 m m = 0,2 kg α = 60° Δt = 0,01 s Asked: The force needed by the golfer to hit the ball so that it goes right into hole C?	3	• Identify the problem according to the concept • Make a list of known quantities • Determine the amount in question
		Diagnosing the problem At the required time B-C = time A-C. On the x-axis, the object moves at a constant speed so that the speed to travel A-C is the same as the speed of the ball after being hit.	3	• Make a diagram, sketch, or symbol that describes the problem • Determine the relationship between quantities
		Planning and implementing plans x = Vo.cosα.t 5 = Vo . cos 60° . 1 5 = Vo . 1/2 Vo = 10 m/s When the ball has not been hit, its velocity = 0, and after being hit its speed = 10 m/s. When a golfer hits the ball: F.Δt = m.ΔV F . 0,01 = 0,2 (10 - 0) F . 0,01 = 2 F = 200 N	3	• Determine the right equation for problem-solving • Substituting the magnitude values into the equation • Perform calculations using the selected equation
		Monitoring and reflecting on solutions So, the force needed by golf players to put the ball into hole C is to be hit by a force of 200 N.	3	• Evaluate the suitability of the solution to the concept • Evaluating units

Figure 1. The item of MoIP-ST

Following the completion of all MoIP-ST items, expert validation was performed on three validators. The validator conducted the assessment, following the three assessment domains. The realm of assessment is divided into three parts: material, construction, and language. These domains are translated into 13 assessment instrument aspects. The assessment aspect consists of (A1) items according to the indicators of problem-solving; (A2) clear boundaries of questions and expected answers; (A3) the material asked is following the purpose of measurement; (A4) the content of the material being asked is according to the level, type of school, and grade level; (A5) sentence formulation in the form of interrogative sentences or orders; (A6) there are clear instructions on how to do the questions; (A7) there are scoring guidelines; (A8) there is a relationship between the picture and the problem being asked; (A9) the item questions do not depend on the previous items; (A10) communicative sentence formulation; (A11) sentences using good and correct language and following the variety of languages; (A12) the formulation of the sentence does not give rise to multiple interpretations; and (A13) using common language and verbs (not the local language). Expert validation results were analyzed using a multi-rater Rasch measurement. Figure 2 illustrates the results of the multi-rater analysis.

Figure 2 is split into five columns. The size column (logit transformation) represents the measurement results with values ranging from +2 (top) to -4 (bottom); this number is referred to as the logit value. The second column describes the distribution of items with logit values ranging from less than logit -2 (I1) to logit +2. Experts consider a logit price of 0 to be a good minimum criterion for item quality. If the value is positive (greater than zero logit), it indicates that the panel of experts thinks the item is good. If the logit value is negative (less than zero), it indicates that the item is not considered good.

```
+------------------------------------------------------------------------------------------------+
|Measr|+item           |-Aspect                                                  |-expert|Scale|
|-----+----------------+---------------------------------------------------------+-------+-----|
|   2 + I3  I5  I6  I7 +                                                         +       + (4) |
|     |                |                                                         |       | --- |
|     | I2             |                                                         |       |     |
|     |                |                                                         |       |     |
|   1 +                +                                                         +       +     |
|     |                |                                                         |       |     |
|     |                |                                                         |       |     |
|     | I4             |                                                         |       |     |
*   0 *                * A2   A3                                                 *       *  3  *
|     |                |                                                         |       |     |
|     |                |                                                         | A     |     |
|     |                |                                                         |       |     |
|  -1 +                +                                                         +       +     |
|     |                |                                                         |       |     |
|     |                |                                                         |       |     |
|     | I1             |                                                         |       | --- |
|  -2 +                +                                                         +       +     |
|     |                |                                                         |       |     |
|     |                |                                                         |       |     |
|     |                |                                                         | B     |     |
|  -3 +                +                                                         +       +     |
|     |                |                                                         |       |     |
|     |                |                                                         |       |     |
|     |                |                                                         | C     |     |
|  -4 +                + A1   A10  A11  A12  A13  A4   A5   A6   A7   A8   A9    +       + (2) |
|-----+----------------+---------------------------------------------------------+-------+-----|
|Measr|+item           |-Aspect                                                  |-expert|Scale|
+------------------------------------------------------------------------------------------------+
```

Figure 2. The result of the multi-rater analysis

Based on Figure 2, the A2 assessment aspect (clear question boundaries and expected answers) and the A3 assessment aspect (the material asked according to the measurement objectives) are the most difficult aspects of the assessment to achieve because several items still do not fulfill this. aspects according to the validator's assessment. Aspects of assessment A1, A10, A11, A12, A13, A4, A5, A6, A7, A8, and A9 are the easiest aspects of assessment because all items match them. According to the expert's opinion, points I3, I5, I6, and I7 are items that fulfill all aspects of the assessment very well. While point I1 does not meet the assessment aspects A2 and A3. This is in accordance with the expert's opinion that the lower the logit value of the assessment aspect, the easier it is for the assessment aspect to fulfill an item. Whereas the higher the logit value of the assessment aspect, the more difficult the assessment aspect is to be fulfilled by an item (Darmana et al., 2021). The expert who gives the most generous value is expert C. In sequential order, the expert who gives the value from the generous to the stingy is expert C, expert B, and expert C. Because expert C has the lowest logit value and expert A has the highest logit value. The assessment of each appraiser can also be analyzed to determine the reliability of the appraiser, as shown in Figure 3.

Total Score	Total Count	Obsvd Average	Fair(M) Average	- Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	Estim. Discrm	Correlation PtMea	Correlation PtExp	Exact Agree. Obs %	Exact Agree. Exp %	N expert
359	91	3.95	3.16	-.58	.77	.99	.1	.99	.1	1.04	.64	.60	96.2	96.2	1 A
362	91	3.98	3.74	-2.72	.98	1.25	.5	1.07	.3	.79	.47	.47	97.8	97.2	2 B
363	91	3.99	3.90	-3.87	1.19	.75	-.1	.36	.0	1.28	.42	.36	97.3	97.0	3 C
361.3	91.0	3.97	3.60	-2.39	.98	.99	.2	.81	.2		.51				Mean (Count: 3)
1.7	.0	.02	.32	1.36	.17	.21	.3	.32	.1		.10				S.D. (Population)
2.1	.0	.02	.39	1.67	.21	.25	.4	.39	.2		.12				S.D. (Sample)

Model, Populn: RMSE .99 Adj (True) S.D. .93 Separation .94 Strata 1.58 Reliability (not inter-rater) .47
Model, Sample: RMSE .99 Adj (True) S.D. 1.34 Separation 1.35 Strata 2.13 Reliability (not inter-rater) .64
Model, Fixed (all same) chi-squared: 6.4 d.f.: 2 significance (probability): .04
Model, Random (normal) chi-squared: 2.1 d.f.: 1 significance (probability): .15
Inter-Rater agreement opportunities: 273 Exact agreements: 265 = 97.1% Expected: 264.3 = 96.8%

Figure 3. Expert measurement report

Figure 3 presents the results of the quality of expert judgment. All experts provide consistent judgments based on statistical status (Boone et al., 2014). Because the Outfit MNSQ and Outfit ZSTD values for members A, B, and C are all in the range of 0.5–1.5 (Outfit MNSQ) and -2 to +2 (Outfit ZSTD). The reliability between experts is low (0.47), indicating that experts give the same score (Koçak, 2020). Low reliability is caused by the tendency of the rater. Appraisers who tend to give the same assessment can cause low assessment reliability (Bond & Fox, 2015). However, the data generated supports the measurement model, as evidenced by the chi-square test value, which produces a significance (probability) close to 0.00. The similarity of the five experts' assessments (exact agreement) reached 97.1%, which shows that the five experts' assessments were almost the same. This is also in line with the previous literature, which states that the tendency of assessors and assessor behavior such as leniency and severity affect the reliability of assessors (Brookhart et al., 2006; Darmana et al., 2021; Güler, 2014). Thus, the assessment aspect of the expert's assessment results can be used to correct point I1 on the MoIP-ST instrument.

Disseminating

Disseminating is the implementation of the developed MoIP-ST instrument. The MoIP-ST was distributed to the students and then the results of the students' work were examined in two stages of analysis. The first analysis is the analysis of the MoIP-ST instrument in terms of validity, reliability, and item difficulty level using the Rasch Model. The second analysis explains students' problem-solving abilities.

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
7	158	64	-.57	.15	2.06	5.00	2.20	5.01	A .34	.66	23.4	46.9	I7
4	74	64	1.24	.16	1.06	.34	.79	-.64	B .78	.82	54.7	59.2	I4
1	133	64	-.06	.14	.87	-.79	.99	.02	C .71	.72	28.1	38.8	I1
2	195	64	-1.55	.19	.73	-1.14	.74	-1.15	D .64	.60	54.7	59.6	I2
3	120	64	.20	.14	.71	-1.84	.72	-1.69	c .84	.75	40.6	36.8	I3
5	125	64	.10	.14	.65	-2.35	.65	-2.28	b .86	.74	42.2	37.9	I5
6	99	64	.64	.15	.63	-2.30	.63	-1.99	a .87	.79	59.4	51.4	I6
MEAN	129.1	64.0	.00	.15	.96	-.44	.96	-.39			43.3	47.2	
P.SD	36.3	.0	.82	.02	.47	2.39	.52	2.32			12.8	9.1	

Figure 4. Output of item fit order

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units

		Eigenvalue	Observed		Expected
Total raw variance in observations	=	20.8400	100.0%		100.0%
Raw variance explained by measures	=	13.8400	66.4%		65.2%
Raw variance explained by persons	=	8.6512	41.5%		40.7%
Raw Variance explained by items	=	5.1888	24.9%		24.4%
Raw unexplained variance (total)	=	7.0000	33.6%	100.0%	34.8%
Unexplned variance in 1st contrast	=	2.4890	11.9%	35.6%	
Unexplned variance in 2nd contrast	=	1.2691	6.1%	18.1%	
Unexplned variance in 3rd contrast	=	1.2266	5.9%	17.5%	
Unexplned variance in 4th contrast	=	.9892	4.7%	14.1%	
Unexplned variance in 5th contrast	=	.5941	2.9%	8.5%	

Figure 5. Unidimensionality of MoIP-ST

Based on Figure 4, the results of the MoIP-ST instrument analysis on item I7 do not meet the Outfit MNSQ criteria because it is outside the value range of 0.5 < MNSQ < 1.5. In addition, item I7 also does not meet the criteria for the outfit values ZSTD and PT MEAN CORR because each is outside the ranges of -2.0 < ZSTD +2.0 and 0.4 < PT MEAN CORR < 0.85 (Sumintono & Widhiarsho, 2015). Items I5 and I6 are retained because they still meet one or two criteria. Thus, only item I7 needs to be removed. However, the validity of the Rasch model was examined not only for each item but also as a whole by examining the results of the instrument's unidimensionality. The instrument's unidimensionality is a criterion for determining whether or not the developed instrument can measure what it is supposed to measure. The results of the unidimensionality test are depicted in Figure 4. The raw variance explained by measure is 66.4% greater than 60%. Based on these findings, the overall validity of the MoIP-ST instrument can be classified as excellent (Mofreh et al., 2014; Ng et al., 2018; Sumintono & Widhiarsho, 2015; Sumintono & Widhiarso, 2015; Talib et al., 2019). Meanwhile, all unexplained variance eigen values in the first have a value of less than 3, and the observed have a value of less than 15% (Mofreh et al., 2014; Ng et al., 2018; Sumintono & Widhiarso, 2015; Talib et al., 2019). Thus, the MoIP-ST instrument used to assess students' problem-solving shows valid results.

In addition to validity analysis, instrument reliability analysis was performed on the MoIP-ST instrument. Following the output summary statistics, a reliability analysis was performed. Figure 5 depicts the output of the summary statistics.

```
     SUMMARY OF 64 MEASURED Person
-------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT          OUTFIT    |
|          SCORE      COUNT     MEASURE   S.E.      MNSQ   ZSTD    MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      14.1        7.0        -.02     .49       .98    .02     .96    .05 |
|  SEM        .8         .0         .19     .02       .06    .11     .07    .10 |
| P.SD       6.7         .0        1.51     .12       .48    .86     .53    .79 |
| S.SD       6.7         .0        1.52     .12       .48    .87     .54    .80 |
| MAX.      26.0        7.0        3.48     .84      3.12   2.82    3.41   2.91 |
| MIN.       6.0        7.0       -1.60     .39       .31  -1.94     .29  -1.79 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .54 TRUE SD    1.41  SEPARATION  2.59  Person RELIABILITY  .87 |
|MODEL RMSE    .50 TRUE SD    1.42  SEPARATION  2.83  Person RELIABILITY  .89 |
| S.E. OF Person MEAN = .19                                                    |
-------------------------------------------------------------------------------
Person RAW SCORE-TO-MEASURE CORRELATION = .98 (approximate due to missing data)
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .87  SEM = 2.45
STANDARDIZED (50 ITEM) RELIABILITY = .98
 |
     SUMMARY OF 7 MEASURED Item
-------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT          OUTFIT    |
|          SCORE      COUNT     MEASURE   S.E.      MNSQ   ZSTD    MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN     129.1       64.0         .00     .15       .96   -.44     .96   -.39 |
|  SEM      14.8         .0         .34     .01       .19    .97     .21    .95 |
| P.SD      36.3         .0         .82     .02       .47   2.39     .52   2.32 |
| S.SD      39.2         .0         .89     .02       .51   2.58     .56   2.51 |
| MAX.     195.0       64.0        1.24     .19      2.06   5.00    2.20   5.01 |
| MIN.      74.0       64.0       -1.55     .14       .63  -2.35     .63  -2.28 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .16 TRUE SD     .80  SEPARATION  4.89  Item   RELIABILITY  .96 |
|MODEL RMSE    .15 TRUE SD     .81  SEPARATION  5.24  Item   RELIABILITY  .96 |
| S.E. OF Item MEAN = .34                                                      |
-------------------------------------------------------------------------------
Item RAW SCORE-TO-MEASURE CORRELATION = -1.00 (approximate due to missing data)
Global statistics: please see Table 44.
UMEAN=.0000 USCALE=1.0000
```

Figure 6. Output of summary statistic

Cronbach Alpha values, person reliability, and item reliability are all part of the Rasch Model reliability test. Figure 6 depicts the results of the Rasch reliability test. According to the MoIP-ST instrument analysis results in Figure 6, item reliability is 0.96 on an excellent interpretation and person reliability is 0.87 on a good interpretation. Cronbach Alpha has a value of 0.87, which is considered very good (Sumintono & Widhiarsho, 2015). The analysis of the MoIP-ST instrument results shows that the test items are of very high quality, and the consistency of the answers given by students is also high. Also noteworthy is the interaction between the students and the items as a whole.

The items' difficulty level can be classified by looking at the output of the item measure in the section the measure logit value and the standard deviation (SD) value (Sumintono & Widhiarsho, 2015). The items' difficulty level is determined by the logit value. The higher the logit value of an item, the more difficult it is. The items in the MoIP-ST are classified as shown in Table 3 below based on the logit measure and standard deviation in Figure 7.

```
Item STATISTICS:  MEASURE ORDER

-------------------------------------------------------------------------------------------
|ENTRY   TOTAL  TOTAL    JMLE   MODEL|   INFIT    |  OUTFIT    |PTMEASUR-AL|EXACT MATCH|      |
|NUMBER  SCORE  COUNT  MEASURE  S.E. |MNSQ  ZSTD|MNSQ  ZSTD|CORR.  EXP.| OBS%  EXP%| Item |
|------------------------------------+----------+----------+-----------+-----------+------|
|    4     74     64    1.24     .16|1.06    .34| .79  -.64| .78    .82| 54.7  59.2| I4   |
|    6     99     64     .64     .15| .63 -2.30| .63 -1.99| .87    .79| 59.4  51.4| I6   |
|    3    120     64     .20     .14| .71 -1.84| .72 -1.69| .84    .75| 40.6  36.8| I3   |
|    5    125     64     .10     .14| .65 -2.35| .65 -2.28| .86    .74| 42.2  37.9| I5   |
|    1    133     64    -.06     .14| .87  -.79| .99    .02| .71    .72| 28.1  38.8| I1   |
|    7    158     64    -.57     .15|2.06  5.00|2.20  5.01| .34    .66| 23.4  46.9| I7   |
|    2    195     64   -1.55     .19| .73 -1.14| .74 -1.15| .64    .60| 54.7  59.6| I2   |
|------------------------------------+----------+----------+-----------+-----------+------|
| MEAN   129.1   64.0     .00     .15| .96  -.44| .96  -.39|           | 43.3  47.2|      |
| P.SD    36.3     .0     .82     .02| .47  2.39| .52  2.32|           | 12.8   9.1|      |
-------------------------------------------------------------------------------------------
```

Figure 7. Output of item measure

Table 3. Item difficulty level of moip-st

Measure Logit	Interpretations of The Difficulty Level	Item	Number Of Items
$M < -0,82$	Very easy	I2	1
$-0,82 \leq M \leq 0$	Easy	I7. I1	2
$0 \leq M \leq 0,82$	Difficult	I5. I3. I6	3
$0,82 > SD$	Very Difficult	I4	1

Table 3 explains that the difficulty level of the MoIP-ST items varies. Items with a difficulty level of "very difficult" were 14.3%, the difficulty level of "difficult" was 42.8%, the difficulty level was "easy" was 28.6%, and the difficulty level was "very easy" was 14.3%. Thus, the results of the analysis of the difficulty level of the items show that the majority of the interpretations of the difficulty level of the MoIP-ST items are "difficult".

Figure 8 illustrates the results of the distribution of students' problem-solving abilities with the MoIP-ST instrument on the left-hand Wright map. Each student is coded according to the serial number. The student code is followed by the student gender code, namely female (F) and male (M). In addition, the student code is also followed by the student code consisting of the urban (U) and rural (R). The student's problem-solving abilities range from a logit value of more than -2 (bottom) to less than +4 (top). Students who have the highest problem-solving abilities are students 10FU, 19FU, and 25MU. The logit value of these students is more than +3 logit. of those showing very high and different problem-solving abilities (outliers). Students 10FU, 19FU, and 25MU are also outside the limits of two standard deviations (T). Students who have the highest problem-solving abilities are dominated by students from cities (C). The lowest problem solving abilities were occupied by students 34FR, 38FR, 39FR, 40FR, 42FR, 51FR, 55FR, 57FR, 60FR, and 64FR. It can be said that these students do not have the problem-solving abilities to answer problems in almost all items. Students who are at the lowest problem-solving ability come from the rural.

When viewed from the highest and lowest problem-solving abilities, it can be concluded that the problem-solving abilities of urban students are better than rural students.

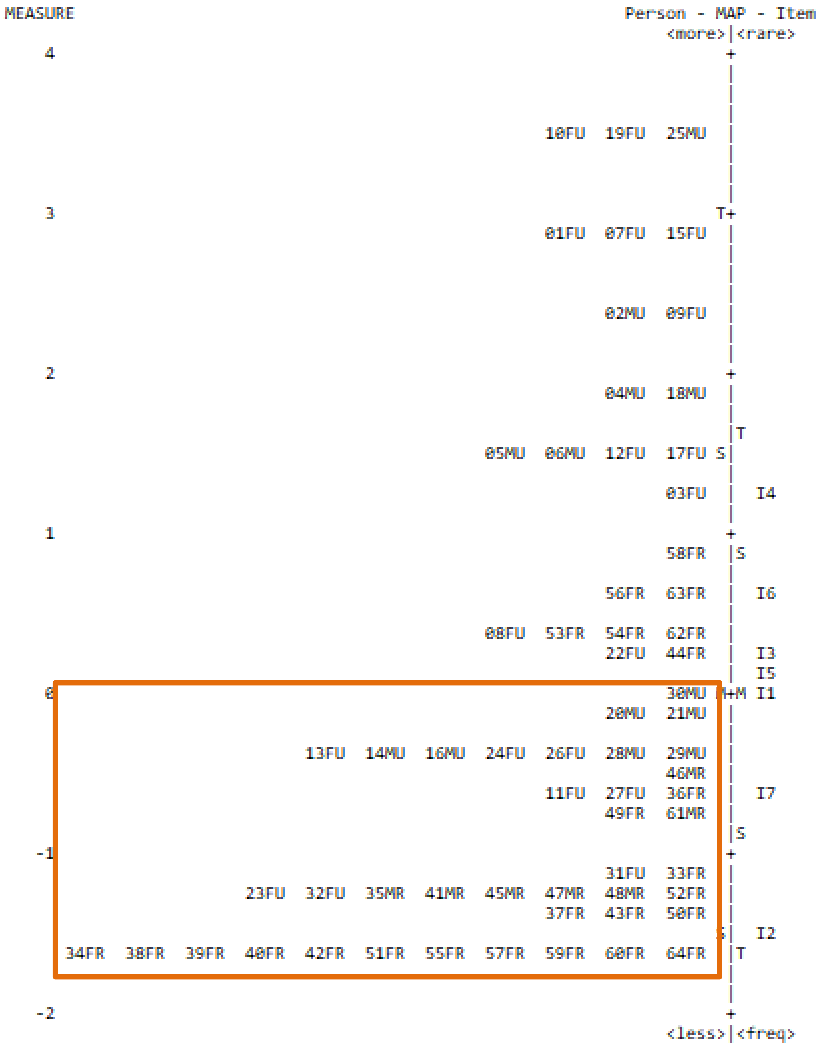


Figure 8. Output wright maps

Students with logit values less than zero are 40, while students with logit values greater than zero are 24. This demonstrates that more students have below-average problem-solving abilities. These results are also consistent with the average logit person's result, which is -0.2 logit, which means it is less than 0.0 logit. This demonstrates that students' problem-solving abilities are below the MoIP-ST item items' average difficulty level. Under 0 logit students include 15 urban students and 25 rural students. These findings indicate that rural students' ability remains low when compared to urban students.

On the topic of momentum and impulse, the average percentage of students achieving problem-solving abilities as a whole score from urban and rural students is 61% and 37%, respectively. The overall value of students' problem-solving skills is in the medium range for both urban and rural students. The average percentage of problem-solving abilities at each stage is obtained when students' problem-solving abilities are broken down at each stage, as shown in Figure 1.

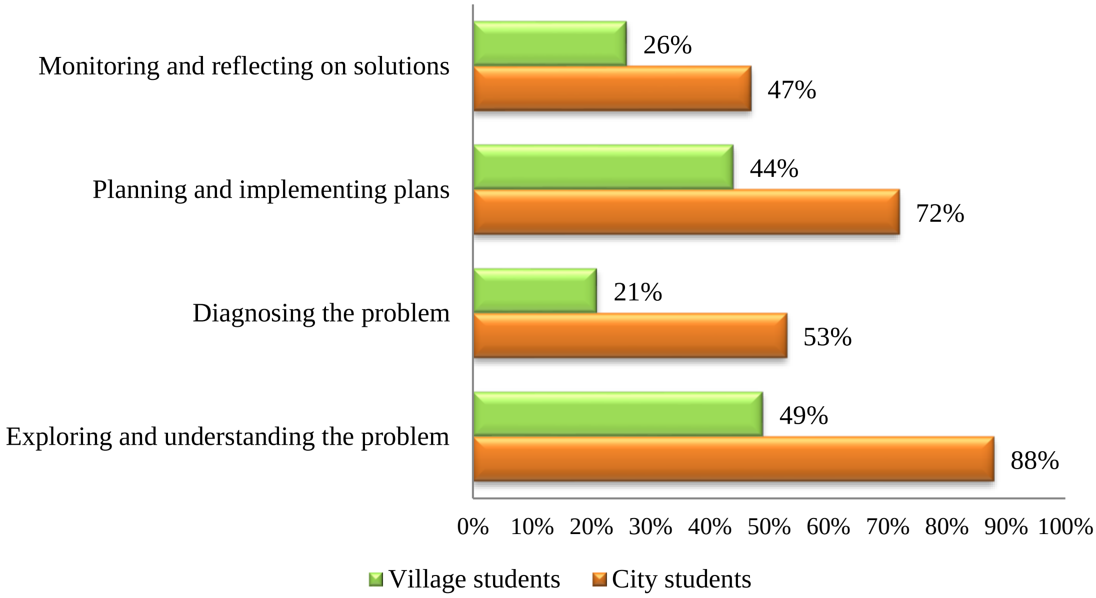


Figure 9. Percentage of problem-solving abilities at each stage

Figure 9 shows that urban students outperform rural students at each stage of problem solving. The ability to solve problems at the exploring and understanding stage is where the majority of urban and rural students excel. The lowest stage for urban students is monitoring and reflecting on solutions, while the lowest stage for rural students is diagnosing the problem. Students struggle to re-present a problem in the form of graphs, tables, pictures, words, and symbols, so the stage of diagnosing the problem receives a low score. The monitoring and reflecting stages are not difficult because you only need to monitor and reflect on the solutions obtained at this stage, but students frequently forget and miss them.

Students' failure at the problem-diagnosis stage demonstrates that they are unable to construct relationships between known and unknown variables. Students are unable to provide useful descriptions to begin problem-solving strategies. Furthermore, students continue to lack proficiency in drawing physical sketches to represent problems. This inability has an impact on their inability to define key concepts or principles, resulting in formulations that do not demonstrate the correct relationship. The formulated physical equations do not correspond to the physics sketch, physics concepts, principles, or the specified specific conditions. Some students can draw physics sketches and determine the correct physics principles but are unable to formulate equations that lead to the expected completion. Some other students are not able to go through the diagnostic stage of the problem correctly but can do the planning and executing stages. This shows that students are still oriented only to mathematics when solving physics problems, not to the right and appropriate concepts and strategies.

The study's low yields are consistent with Diawati's findings (Diawati, 2016). Conceptual understanding can influence students' problem-solving abilities (Sutarno et al., 2017). Problems with students' perceptions can lead to students' inability to solve problems. Students' perceptions can also confirm students' lack of problem-solving abilities (Diawati, 2016). Students who have an understanding of physics concepts and can store physics principles in their memory as inseparable information, tend to be better able to solve the problems they face. Students' skills in solving problems are also influenced by their knowledge and previous experience solving problems. Learner experiences can be formed from models, methods, and strategies for learning in the classroom and the laboratory. So those students are not accustomed to using the steps of problem-solving strategies, practice questions are not oriented to non-routine problems, and the learning strategies applied have not trained problem-solving skills (Sutarno et al., 2017). In addition, students' problem-solving abilities that are lacking can be influenced by a lack of independent teaching materials to help students during the learning process. Worksheets is one of the teaching materials in the form of sheets that are useful for students as learning instructions and contain assignments in the form of questions and practicum activities that must be completed by students (Fadilah & Yohandri, 2019). Worksheets can help students understand physics problems related to everyday problems. Giving worksheets to students can strengthen students who complete assignments and provide guidance to students who have difficulty mastering problems so that students can develop students mindsets (Fadilah & Yohandri, 2019). The use of worksheets in learning has the potential to train students' problem-solving abilities to become better at solving problems (Wulantri et al., 2020).

▪ CONCLUSION

The result of this research is that the MoIP-ST instrument that has been developed can have good quality in terms of reliability, but in terms of item validity in item I7, it needs to be improved. The difficulty of the items in MoIP-ST varies. The MoIP-ST instrument is appropriate for use in class during learning because of its validity, reliability, and level of difficulty. The results of using the MoIP-ST on problem-solving abilities show that students' problem-solving abilities need to be improved through models, methods, and learning models that support problem-solving. As well as the creation of appropriate problem-solving worksheets.

▪ REFERENCES

Baig, M., Ali, S. K., Ali, S., & Huda, N. (1969). Quality evaluation of assessment tools: OSPE, SEQ & MCQ. *Pakistan Journal of Medical Sciences*, *30*(1), 3–6.

Bond, T. G., & Fox, C. M. (2015). Applying the Rasch Model: Fundamental Measurement in the Human Sciences, Third Edition 3rd Edition. In *Maternal & Child Nutrition* (Vol. 11, Issue supplement S2).

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.

https://doi.org/10.1207/s15324818ame1902_5

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*(2), 121–152.

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, G. A. C., Sunarno, W., & Supriyanto, A. (2019). The needs analysis on module development based on creative problem solving method to improve students' problem solving ability. *Journal of Physics: Conference Series*, *1153*(1).

Diawati, C. (2016). Students' conceptions and problem-solving ability on topic chemical thermodynamics. *AIP Conference Proceedings*, *1708*(February 2016). https://doi.org/10.1063/1.4941152

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), 1–13.

Fadilah, R., & Yohandri. (2019). Need analysis of student worksheet based on tracker on static fluid learning material in high school. *Journal of Physics: Conference Series*, *1185*(1).

Gok, T. (2010). The General Assessment of Problem Solving Processes and Metacognition in Physics Education. *International Journal of Physics & Chemistry Education*, *2*(2), 110–122.

Güler, N. (2014). Analysis of Open-Ended Statistics Questions with Many Facet Rasch Model. *Eurasian Journal of Educational Research*, *55*, 73–90. https://doi.org/10.14689/ejer.2014.55.5

Haladyna, T. M., & Rodriguez, M. C. (2013). Developing and validating test items. *Developing and Validating Test Items*, 1–446. https://doi.org/10.4324/9780203850381

Hegde, B., & Meera, B. N. (2012). How do they solve it? An insight into the learner's approach to the mechanism of physics problem solving. *Physical Review Special Topics - Physics Education Research*, *8*(1), 1–9.

Henderson, C. R., Heller, K., Heller, P., Kuo, H. V., & Yerushalmi, E. (2016). *Instructors' Ideas about Problem Solving – Setting Goals*. *November 2015*.

J. Sirait, L. Sutrisno, N. Balta, & A. Mason. (2017). the Development of Questionnaire To Investigate Students' Attitudes and Approaches in Physics Problem Solving. *Jurnal Pendidikan Fisika Indonesia*, *13*(2), 79–87. http://journal.unnes.ac.id/nju/index.php/jpfi

Jonassen, D. H. (2011). *Learning to Solve Problems*.

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

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.

Kohl, P. B., & Finkelstein, N. D. (2008). Patterns of multipe representation use by experts and novices during physics problem solving. *Physical Review Special Topics - Physics Education Research*, *4*(1), 1–13.

Kubiszyn, T., & Borich, G. D. (2013). *Educational Testing and Measurement: Classroom Application and Practice, 10th Edition*. NJ: Willey. https://books.google.cz/books?id=0VkcAAAAQBAJ

Masitoh, P. N. A., Latifah, S., Saregar, A., Aziz, A., Suharto, & Jamaluddin, W. (2021). Bibliometric analysis of physics problem solving. *IOP Conference Series: Earth and Environmental Science*, *1796*(1).

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

Nadapdap, A. T. Y., & Istiyono, E. (2017). Developing physics problem-solving skill test for grade X students of senior high school. *Research and Evaluation in Education*, *3(2)*(1), 114–123.

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. (2014). *PISA 2012 Results: Creative Problem Solving: Students' Skills in Tackling Real-Life Problems (Volume V): Vol. V*. OECD Publishing.

Polya, G. (1957). "How to Solve It." In *How to Solve It*.

Schoenfeld, A. (1992). Handbook of research on mathematics teaching and learning. In *Choice Reviews Online* (Vol. 30, Issue 04).

Sujarwanto, E., Hidayat, A., & Wartono. (2014). Jurnal Pendidikan IPA Indonesia. *Jurnal Pendidikan IPA Indonesia*, *3*(2), 146–153.

Sumintono, B., & Widhiarsho, W. (2015). *Aplikasi Pemodelan Rasch pada Assesment Pendidikan*. Trim Komunikata.

Sumintono, B., & Widhiarso, W. (2015). Penilaian Pendidikan dan Ujian. *AplikAsi RascH PemodelAn Pada Assessment Pendidikan*, 1–4.

Sutarno, S., Setiawan, A., Kaniawati, I., & Suhandi, A. (2017). Pre-Service Physics Teachers' Problem-solving Skills in Projectile Motion Concept. *Journal of Physics: Conference Series*, *895*(1).

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 assessment Survey (CETAS). *Universal Journal of Educational Research*, *7*(12 A), 10–17. https://doi.org/10.13189/ujer.2019.071902

Wulantri, Distrik, I. W., Suyatna, A., & Rosidin, U. (2020). The Effectiveness of Creative-Inquiry-Based Student Worksheet in Improving Physics Self-Efficacy and Problem Solving of Senior High School Students. *Journal of Physics: Conference Series*, *1467*(1).