Journal for the Education of Gifted Young Scientists, 7(4), 1037-1053, December 2019
e-ISSN: 2149- 360X
http://jegys.org

Research Article

The Development of High School Physics Problem Solving Skills Test Instruments Based Problem-Based Learning

Ahsan ABDULFATTAH[1], SUPAHAR[2]

Received: 6 August 2019 **Accepted:** 3 December 2019

Abstract
Problem solving skills are an important part of physics learning in schools and is also useful for adapting to the environment. Problem solving skills are also a demand for education in the 21st century, so having this skills can help to compete in gaining experience in this all-modern world. This study aimed to determine the feasibility of the developed test instrument in terms of content validity and reliability. The test instrument developed was in the form of multiple choices with a total of 25 questions. The analysis of content validity of the test instrument was conducted using the Aiken's V Coefficient. The empirical validity and reliability of test instrument were estimated using the *Classical Test Theory* (CTT) and *Item Response Theory* (IRT). The results of this study showed that, the test instrument developed was feasible to be used as an instrument for testing problem solving skills with valid result according to Aiken's V Coefficient ranging between 0.76 to 0.95 and reliable with a reliability value according to CTT of 9.5, while the developed questions were stated reliable according to the IRT if used by the students with the ability ranging between -1.2 to 3.8 in logit scale.
Keywords:
validity, reliability, test instrument, problem solving skills

To cite this article:

Abdulfattah, A., Supahar. (2019). The development of problem-based learning test instruments for the high school physics problem solving skills. *Journal for the Education of Gifted Young Scientists, 7*(4), 1037-1053. DOI: http://dx.doi.org/10.17478/jegys.602291

[1] Department of Physics Education, Graduate Program, Universitas Negeri Yogyakarta. It is abdulfattahahsan@gmail.com. Orcid no: 0000-0001-8951-2202
[2] Dr., Department of Physics Education, Graduate Program, Universitas Negeri Yogyakarta, Physics Education Lecturer. It is supahar@uny.ac.id . Orcid no: 0000-0002-2486-5549

Introduction

Education has become a basic need in this modern age. Without education, people will find some difficulties in following the development of science and technology as well as difficulties in getting jobs. In order to realize a good education, the right learning process is needed in every educational activity. The learning process includes three aspects, namely learning methods (Lestari, Saryantono, Syazali, Saregar, Madiyo, Jauhayiyah, Umam, 2019), learning media (Huda, Sholikhakh, Bina, Lestari, Habibi,Suharso, (2019), and assessment instruments (Hadarah, Tulhikmah, 2019). Based on the current educational curriculum in Indonesia, the 2013 Curriculum, one of the learning models that can be used, according to the Regulation of the Minister of Education and Culture No. 22 of 2016, is the *Problem-Based Learning* (PBL) model.

According to Asrofi (2008), the implementation of effective and interesting learning requires learning models with the help of appropriate media in order to support learning activities, so that the students will be happy in participating in learning and can receive the knowledge delivered by the teacher. Based on observations conducted in several Senior High Schools in the Special Region of Yogyakarta, it was found that the interest of students in physics learning was quite high, but they were getting bored as the response to the learning process of physics. The teachers have been actually using the 2016-revised 2013 curriculum in delivering the materials, but the learning methods used by teachers were still limited to the lecturing and demonstrating methods (Basheer, Hugerat, Kortam, Hofstein, 2017). The results of the observation also showed that the assessment carried out by the teacher was still in the form of a written test to measure the ability of *Lower-Order Thinking Skills* (LOTS), so that the results of students' problem solving skills were still classified as low (Obenchain, Ives, 2006). In addition, the teachers have also never used technology such as the use of smartphones in learning (Kim, Ilon, Altmann, 2013).

Therefore, in this study, a test instrument for problem solving skills was developed using PBL model with the Edmodo application, so that it was able to be operated through the smartphone. Edmodo is a free and safe educational learning network to use and provides a simple way for teachers to create and manage online classroom communities and allows students to connect and work with their teachers anywhere and anytime (Balasubramanian, Jaykumar, Fukey, 2014). So that, Edmodo can be used to facilitate learning students both in and out of the classroom.

The development of this test instrument was used for assessing the problem solving skills of the students (Neubert, Mainert, Kretzschmar, Greiff, 2015), as can be seen in Table 1 below,

Table 1.

Operational Framework for the Development of Test Instruments

Conceptual Definition	Conceptual Framework	Operasional Framework
Problem solving skills are seen as a fundamental part of learning science at school (Loucks, 2007). Problem solving skills not only important in school and for work, but also to adapt to the environment (Hwang, Wu, Chen, 2012). Physics problem solving skills can provide effective understanding of physics by connecting various ideas and concepts in various situations encountered in solving a problem (Hedge, Meera, 2012).	Problem solving skills have several indicators according to experts. Problem solving skills indicators according to Ruseffendi (1991) are present the problem more clearly, states the problem in operational form, compile alternative hypotheses and work procedures used, test the hypothesis and do work to get results, and re-check whether it is correct. Polya (1973) mentioning problem solving indicator are understand the problem, planning a settlement, solve the problem, and check again. According to Gunawan (2017) there are five steps in solving physics problems, namely identify and define problems, determine goals and objectives, produce a solution, make an action plan, follow-up action plan.	Problem solving skills are one of the skills needed to be able to compete in this advanced world. Based on the results of the synthesis and adjusting to the momentum impulse material in physics, the instruments test that developed in this study use the following problem solving skills indicators, understanding problems, planning solutions, solving problems, and rechecking.

Research Problem

What is the feasibility and reliability of the 25 items of the multiple choice test instrument problem solving developed in measuring high school physics problem solving skills? The instrument developed is expected to meet valid and reliable criteria.

Method

The valid and reliable test instruments can be used to measure improvements in problem solving skills. The stages of study were presented in Figure 1 below,

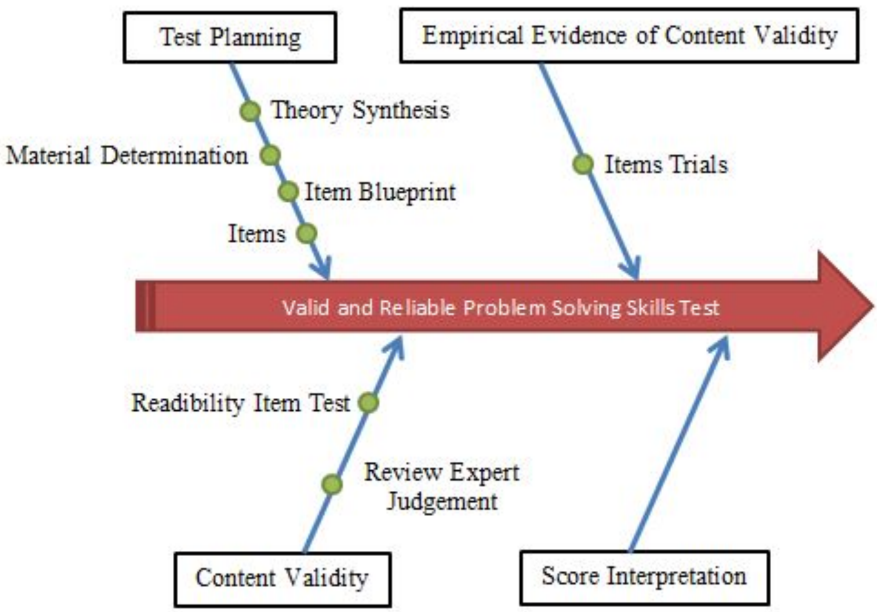

Figure 1.
Development Stages Of Problem Solving Skills Test Instrument

The test instrument developed referred to the indicators of problem solving skills according to Polya (1973), namely understanding problems, planning solutions, solving problems, and re-checking. The indicators were described as an indicators of learning competency obtained from the selected physics materials. Then, the questions draft was made and developed into questions, and the test instrument was given to experts for content validation. The test instrument consisted of 25 questions with a category ranged between C4 to C6. Then, the 25 questions were divided into two package of questions coded A and B, where each package was prepared by considering the representation of each of the measured indicators of problem solving skills. Each package consisted of 15 questions, including 5 questions as the *anchor*.

Participants

The test instrument was put in a trial in January 2019 involving 332 students in four high schools representing high, medium and low category schools based on the results of the 2017 National Science Examination in the Special Region of Yogyakarta. The four schools included SMA Negeri 5 Yogyakarta, SMA Negeri 1 Ngemplak, SMA Negeri 2 Ngaglik, and SMA Negeri 1 Piyungan.

Data Collection

The data obtained in this study were the results of content validation through expert judgment, empirical validation through the results of the trial, and reliability of the test. The content validity of the instrument was obtained by submitting the developed instruments to experts for some review. There were three experts in this study, two physics education evaluation practitioner, and two colleagues. Furthermore, the reviewed test instrument was put into a readability test with 10 high school students as the respondents. The results of expert validation and readability test were used to improve the test instrument.

Data Analysis

The content validity value of the instrument was presented by Aiken (1985) who formulated the Aiken's V formula to calculate the content validity coefficient based on the results of evaluations by a number of experts as many as *n*-people on the extent to which an item can represent the measured construct. The construct referred to here is the relevance of the item with the operational translation of the measured attribute. The Aiken's V statistic is formulated as follow:

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

information:

s = $r-lo$

lo = lowest validity rating

c = highest validity rating

r = the number that given by reviewer

The determination of categories of instrument validity referred to the categories proposed by Azwar (2017), as shown in Table 2,

Table 2.

V Value Criteria

V Value	Categorization
> 0,35	Very Useful
0,21 – 0,35	Useful
0,11 – 0,20	Depends of Condition
< 0,11	Useless

The evidence of the validity of the empirical scores from the trial of the questions was then analyzed using a *Partial Credit Model* (PCM) which is the development of the IRT 1 *Parameter Logistics* (1-PL) model. The data analysis included testing on goodness of fit of the questions with PCM model, determining the index reliability of test, estimating the level of difficulty of questions, estimating the ability parameters, and describing the total information function curves and conducting *Standard Error Measurement* (SEM) (Supahar, 2015). The results of the trial data were analyzed using the QUEST program to find out whether the items developed were fit for the PCM model or not. The rules of Adam and Khoo (1996: 30) state that an item is fit to the PCM model if the mean INFIT MNSQ along with its standard deviation and INFIT value t along with the standard deviation meet the criteria in Table 3 below,

Table 3.
Item Fit Category with PCM Models

No	Parameter	Range
1	INFIT MNSQ	0,77 – 1,30
2	INFIT t	-2,0 – 2,0

The MNSQ INFIT value limits the distribution of the calibrated score and is still on the *Leptokurtic* curve which is still in a state of *unity*.

The QUEST program can also be used to determine the estimated reliability of the test instrument. The requirement for a good reliability estimation follow the opinion of Sumadi Suryabrata (2002), where a test instrument can be used if it has at least a reliability coefficient of 0.90. George and Mallery (2003) categorizes reliability value as shown in Table 4 below,

Table 4.
Reliability Category

Score Range	Category
X ≥ 0,9	Excellent
0,9 > X ≥ 0,8	Good
0,8 > X ≥ 0,7	Acceptable
0,7 > X ≥ 0,6	Questionable
0,6 > X ≥ 0,5	Poor
≤ 0,5	Unacceptable

The reliability values can also be obtained through analysis using the PARSCALE program, which is seen from the *Total Information Curve* (TIC).

Results

Theoretical study is conducted to find out what indicators exist in problem solving skills. Some experts have presented indicators of problem solving skills. This study referred to the skill indicators by Polya, namely understanding problems, planning solutions, solving problems, and re-checking.

The theoretical study is also conducted to determine the materials used. The development of this test instrument was adjusted to momentum and impulses materials. Then, from the achievement indicators of learning competency, consisting of ten indicators, namely: defining the definition of momentum, explaining the concept of impulses, explaining the relationship of the concepts of momentum and impulses, formulating the concepts of momentum and impulses and the interrelations between the two concepts outside, solving problems related to the principles of conservation of momentum, explaining types of collisions, integrating energy conservation laws and conservation of momentum for various collisions, describing the application of momentum conservation laws for systems without external forces with a *bekel* ball game collision experiments, and describing momentum and impulse relations in the *pletokan* (one of blowgun types) game. The draft of the developed test instrument developed is presented in the Table 5 below,

Table 5.

Problem Solving Skills Test Instrument Blueprint

No	Problem Solving Skills Indicator	Basic Competencies Achievement Indicator	Item Indicator
1	Understanding problems	Analyze momentum conservation laws for systems without external forces.	Find a way to increase the speed of a bouncing ball
		Explain impulse concept	Choose quantities that can explain the concept of impulses from pletokan game
		Defines momentum	Interpreting picture illustrations related to momentum concept
		Formulate momentum and impulses concepts, and the interrelationships	Examine the true statement to momentum and impulse concepts
		Integrating energy conservation laws and	Analyze the speed of two objects after collision

		momentum conservation laws for various collision events	Analyze the magnitude of the kinetic energy of two colliding objects
2	Planning solutions	Investigate the application of the momentum conservation law for systems without external forces with bekel balls experiments.	Design an experimental step to obtain data on bekel balls collision experiment
		Explain impulse concept	Focusing on ways to determine the impulse force on pletokan bullet
		Investigate momentum and impulse concept in pletokan game	Design experimental steps in pletokan games to obtain observational data
		Defines momentum	Look at the concept of momentum in a ball that is thrown at a certain speed, and has a certain mass
		Explain collisions types	Examine the causes of perfect elastic collisions on tennis balls and iron balls
3	Solve the problems	Solve problems related to momentum conservation principle	Analyze distance between two objects that work on each other's force by applying momentum conservation laws
			Analyze the boat speed after the last passenger jumps out
		Explain the relationship between momentum and impulse concept	Choose the influence of the impulsive force received by the passenger from the safety equipment on the car when a collision occurs
			Analyze the influence of impulsive forces on collision objects

		Analyze momentum conservation laws for systems without external forces	Analyze speed that people received after firing their arrows
		Integrating energy conservation laws and momentum conservation laws for various collision events	Analyzes the speed of two objects that move together after a collision
			Analyze the speed of two objects after collision
			Analyze the speed of objects after collisions
			Analyze the ratio of kinetic energy in object that collisions
4	Re-checking	Investigate the application of the momentum conservation law for systems without external forces with bekel balls experiments.	Evaluate the results of implementing momentum conservation law with bekel ball collision experiment.
		Explain impulse concept	Review the impulses concept that obtained from pletokan games
		Investigate momentum and impulse concept in pletokan game	Test the results of pletokan experiments that has been done
		Explain collisions types	Choose the cause of the perfect inelastic collision between two Tamiya cars

The test instrument for problem solving skills was validated by seven judgement experts before being used for instrument testing. Validation results were analyzed quantitatively. The results of the test instrument analysis from seven validators showed that the Aiken's coefficient value ranged between 0.76 to 0.95, so all items in this instrument were declared valid and very useful as referred to Table 2. The results of the evaluation by seven validators can be seen in Figure 2 . The results of readability test on the items/questions involving 10 students also showed good results because there were no items that needed to be revised. Therefore, this test instrument is feasible to use.

The Empirical evidence of content validity was obtained by testing the questions that have met the validity according to judgement. There were 25 items stated to meet validity. They were then divided into two packages (Package A and Package B). Each package consisted of fifteen items with five items as anchors (there were five items in package A and B). The trial involved 332 high school students in the Special Region of Yogyakarta. The results were analyzed using the Rasch's IRT model by using the QUEST and Parscale programs.

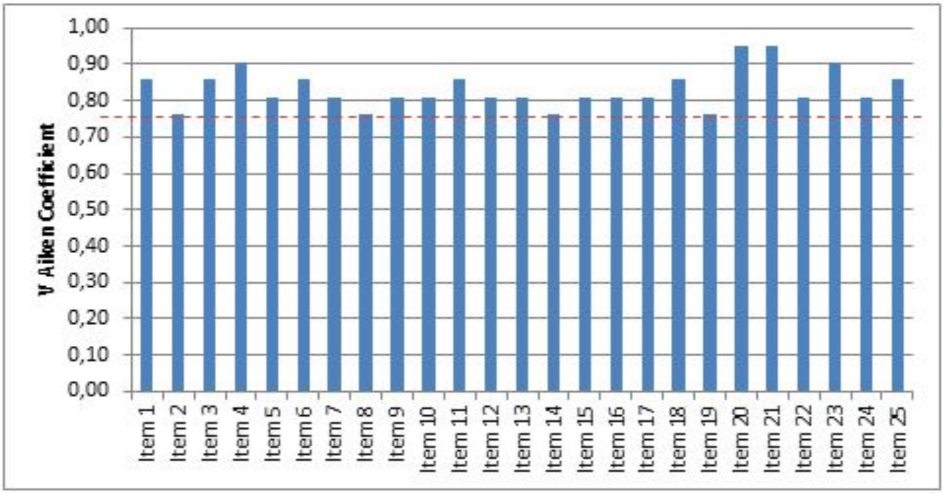

Figure 2.

Expert Validation Result

The results of the trial were in the form of dichotomous 2 data categories which were analyzed according to the Rasch's model. The results of the test match analysis seen from the INFIT parameter for Mean Square (MNSQ) and INFIT t showed that the test instrument of physics problem solving skills meet the *fit statistic* criteria according to the Rasch's model which are completely presented in Table 6. The results of the analysis of 25 items developed have an INFIT MNSQ value ranging between 0.77 to 1.30, so that all test items fit the Rasch's model. The suitability map of the items with the Rasch's model is presented in Figure 3.

Table 6.

Fit Statistics Test Parameter at 0,5 Opportunity Level

No	Test Parameter	Items Estimation	Cases Estimation
1	INFIT MNSQ	1,00 ± 0,04	1,00 ± 0,16
2	OUTFIT MNSQ	0,99 ± 0,17	0,99 ± 0,37
3	INFIT t	0,00 ± 0,78	0,10 ± 0,51
4	OUTFIT t	0,00 ± 1,04	0,12 ± 0,60
5	Mean Difficulty	0,00 ± 0,81	
6	Estimate Reliability	0,95	

The items have a *difficulty index* ranges between -1.87 (items 11) to +1.40 (item 19), so the average *difficulty index* of the test instrument is 0.00 ± 0.81.

```
QUEST: The Interactive Test Analysis System
------------------------------------------------------------------------------------------
Item Fit                                                                    21/ 1/19 13: 4
all on all (N = 332 L = 25 Probability Level= .50)
------------------------------------------------------------------------------------------
INFIT
 MNSQ   .56       .63       .71       .83      1.00      1.20      1.40      1.60      1.8
---------+---------+---------+---------+---------+---------+---------+---------+---------+
  1 item 1                        .             *|              .
  2 item 2                        .              |  *           .
  3 item 3                        .              | *            .
  4 item 4                        .           *  |              .
  5 item 5                        .              |*             .
  6 item 6                        .             *|              .
  7 item 7                        .              *              .
  8 item 8                        .             *|              .
  9 item 9                        .              |    *         .
 10 item 10                       .              *              .
 11 item 11                       .              | *            .
 12 item 12                       .        *     |              .
 13 item 13                       .              |*             .
 14 item 14                       .              *              .
 15 item 15                       .              |   *          .
 16 item 16                       .           *  |              .
 17 item 17                       .            * |              .
 18 item 18                       .             *|              .
 19 item 19                       .              *              .
 20 item 20                       .              |*             .
 21 item 21                       .              |*             .
 22 item 22                       .             *|              .
 23 item 23                       .              *              .
 24 item 24                       .             *|              .
 25 item 25                       .            * |              .
==========================================================================================
```

Figure 3.
INFIT MNSQ Distribution Value

The item characteristics can be seen from the Item Characteristic Curve (ICC) and Item Information Curve (IIC). Figure 4 shows an example of an ICC result for item number 6. The higher the curve of an item, the more valid the item is to be used in the measurement. Item number 6 shows that the item is feasible to use because it has a high peak.

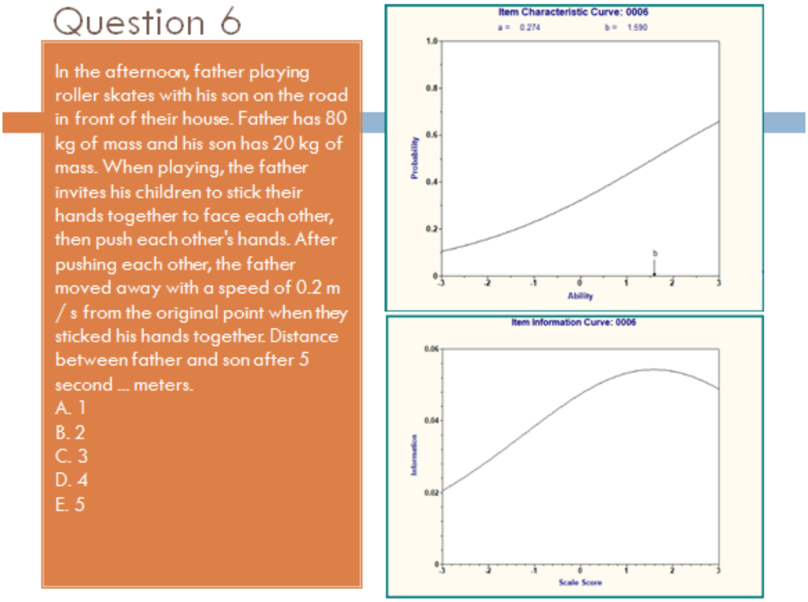

Figure 4.

Item Number 6 Characteristics

Reliability value according to item estimation (sample reliability) were obtained at 9.5 and viewed through the TIC graph in Parscale program obtained that the developed questions were reliable if used by the students with the ability ranging between -1.2 to 3.8 in logit scale, as shown in Figure 5.

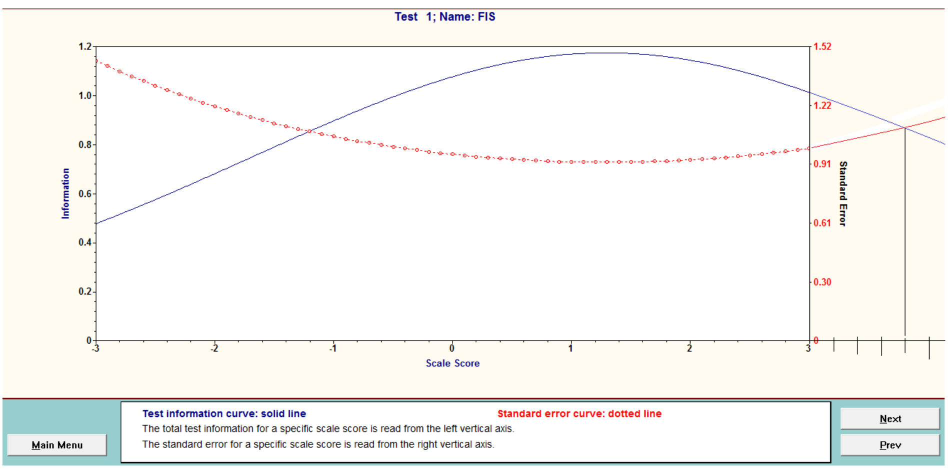

Figure 5.

Total Information Curve (TIC) Test Instrument

The results of difficulty item analysis conducted using the Parscale program also obtained the difficulty level of the item as shown in Figure 6.

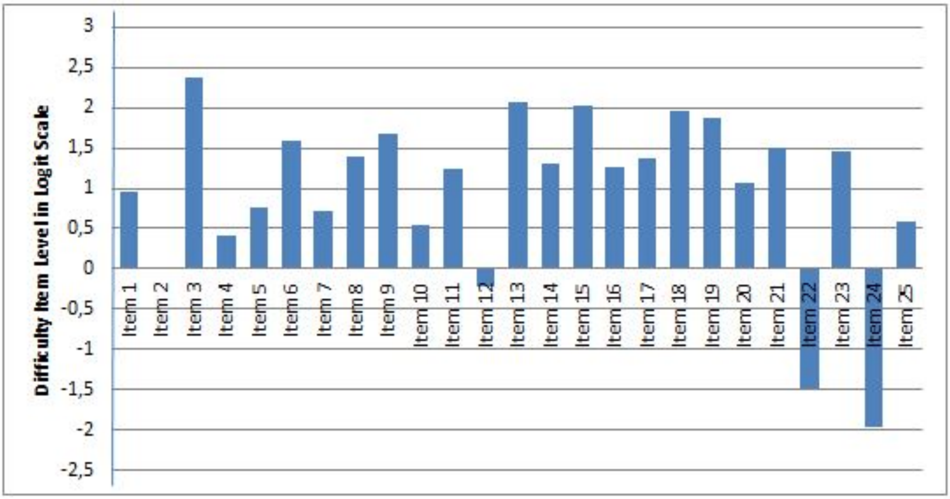

Figure 6.
Difficulty Item Level

The estimation of students' problem solving skills can be known through the *PH3 file* output in the *ability* column that used the logit scale. The value of students' skills is θ (*tetha*). The value of θ in the trial is in the range of between -2.3236 to 3.3855. The students' problem solving skills were then interpreted in four scales according to Djemari Mardapi (2012) as very high, high, low, and very low. The students' problem solving skills are shown in Figure 7.

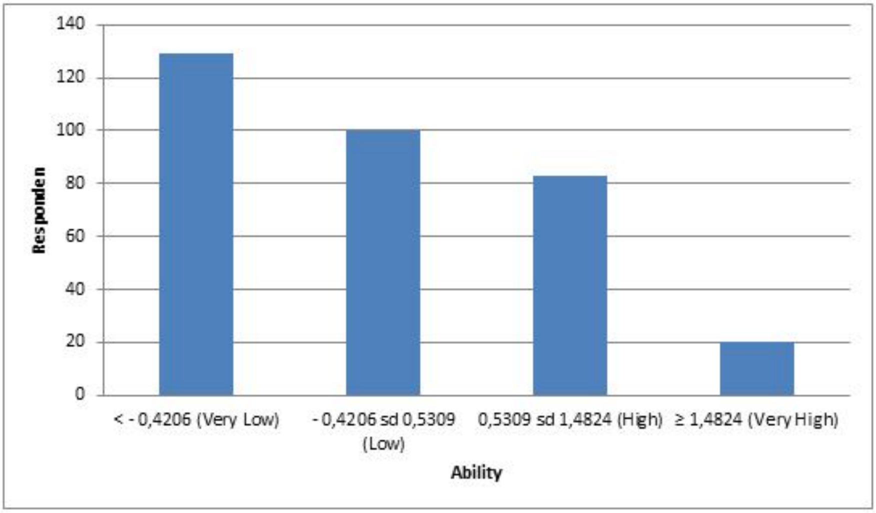

Figure 7.
Students Problem Solving Skills Ability

Discussion and Conclusion

Based on the Aiken's coefficient value, if there were seven validators with four choices of judgment decision, then the coefficient value of V obtained was at least 0.76, so that the items were categorized as valid according to the judgement (Aiken, 1985). Figure 2 shows that the Aiken's V coefficient value obtained ranges between 0.76 to 0.95; thus, according to the judgement, the items developed meet the validity.

The results of the analysis of 25 items developed have an INFIT MNSQ value ranging between 0.77 to 1.30, so that all test items fit the Rasch's model (Adam and Khoo, 1996). The suitability map of the items with the Rasch's model is presented in Figure 3. Reliability value according to item estimation (sample reliability) were obtained at 9.5 with excellent reliability (George and Mallery, 2003) and viewed through the TIC graph obtained that the developed questions were reliable if used by the students with the ability ranging between -1.2 to 3.8 in logit scale, as shown in Figure 5.

The higher difficulty level, the more difficult the item. Conversely, the lower value of the difficulty level of the item, the easier item (Subali & Suyata, 2012). The difficulty level of the items developed is in the range between -1.97 to 2.37 with an average of 0.98 and a standard deviation of 1.04. Item number three has the highest difficulty level and item number 24 has the lowest difficulty level. An item is said to be good if the difficulty level of the item is in the range between -2.0 and 2.0 (-2.0 <b <2.0) (Hambleton & Swaminathan, 1985, p.36). Therefore, based on the difficulty analysis on 25 items, there were 24 good items and 1 bad item because it has a difficulty level of 2.37, which is item number 3.

Based on the results of content validation, the test instrument for problem solving skills was feasible to use which shows that the test instrument was valid with Aiken's V coefficient value ranging between 0.76 to 0.95 and all items were in accordance with the Rasch's model INFIT MNSQ and INFIT t values. Based on the TIC chart, the test instrument is reliable if it is used on students with skills ranging between -1.2 to 3.8 on a logit scale. Thus, the test instrument developed is feasible because it meets the requirements of content, empirical, and reliability validity.

The students' problem solving skills were then interpreted in four scales according to Djemari Mardapi (2012) as very high, high, low, and very low. The students' problem solving skills are shown in Figure 7. There were 6% students with high problem solving skills and 25% students with high problem solving skills. Then, there were 30% students with low problem solving skills, and 39% students with very low problem solving skills.

Biodata of the Author

Ahsan ABDULFATTAH was born in Sukoharjo, Indonesia. He graduated from Department of Physics Education, Faculty of Mathematics and Science of Universitas Negeri Yogyakarta in 2016. He still study to complete his master in the same university's, master of physics education, Universitas Negeri Yogyakarta.

Affiliation:
E-mail: abdulfattahahsan@gmail.com
Phone: 085642001091 / 082331319126

Dr. SUPAHAR was born in Kulon Progo, Indonesia. He graduated from Department of Physics Education IKIP Yogyakarta in 1992. He completed master Department of Physics, ITB Bandung in 2000. And then He completed her doctorate in the Department of Measurement and Assessment Physics Education, Universitas Negeri Yogyakarta in 2014.
Affiliation: Universitas Negeri Yogyakarta, Indonesia.
E-mail: supahar@uny.ac.id **Phone:** 08164267452

References

Adams, R. J. & Khoo, S. T. (1996). *Quest: The interactive test analysis system version 2.1*. Victoria: The Australian Council for Educational Research.

Aiken, L. R.. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement, 45.*

Arends, R.I. (2012). *Learning to Teach*. (9th ed). New York: The McGrow-Hill Companies. Inc.

Asrofi, M.. 2008. Model Pembelajaran yang Menyenangkan. Retrieved from http://www.fisikanet.lipi.go.id.

Azwar, S. (2017). *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar.

Balasubramanian K., Jaykumar V., Fukey L.N. (2014). A study on "Student preference towards the use of Edmodo as a learning platform to create responsible learning environment. *Procedia- Social and Behavioral Sciences, 144,* 416 – 422, https://doi.org/10.1016/j.sbspro.2014.07.311

Basheer, A., Hugerat, M., Kortam, N., Hofstein, A. (2017). The Effectiveness of Teachers' Use of Demonstrations for Enhancing Students' Understanding of and Attitudes to Learning the Oxidation-Reduction Concept. *Eurasia Journal of Mathematics, Science and Technology Education, 13*(3), 555-570. https://doi.org/10.12973/eurasia.2017.00632a

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Boston: Allyn & Bacon.

Gunawan, A. Harjono, H. Sahidu1, L. Herayanti. Virtual Laboratory To Improve Students' Problem-Solving Skills On Electricity Concept. *Jurnal Pendidikan IPA Indonesia, 6*(2), 257-264. https://doi.org/10.15294/jpii.v6i1.8750

Hambleton & Swaminathan. (1985). Item Response Theory Principles and Applications. Boston: Kluwer Nijhoff Publishing.

Hadarah, H, Tulhikmah, R . (2019). Epistemology Philosophy Development of Assessment Instruments Student Chemistry Learning Outcomes. *Journal for the Education of Gifted Young Scientists, 7* (3), 627-641. DOI: 10.17478/jegys.605097

Hedge B., Meera B. N.. (2012). How do they solve it? An insight into the leaner's approach to the mechanism of physics problem solving. *Phisical Review Special Topics – Physics Education Research,* 1-9. https://doi.org/10.1103/PhysRevSTPER.8.010109

Huda, S , Sholikhakh, R , Bina, N , Lestari, F , Habibi, B , Suharso, P . (2019). Effect of Application Smart Circuit Learning Media to Mathematics Learning Outcomes: A Case Study of Islamic School Students. *Journal for the Education of Gifted Young Scientists, 7*(3), 699-715 . DOI: 10.17478/jegys.597053

Hwang, G.J., Wu, P.H., &; Chen, C.C. (2012). An Online Game Approach for Improving Students' Learning Performance in Web-Based Problem Solving Activities. *Computers & Education, 59*(4), 1246–1256. Retrieved from: https://www.learntechlib.org/p/50289/.

Khuluqo, I., E. (2016). Belajar dan Pembelajaran. Jakarta: Pustaka Pelajar.

Kim, J., Ilon, L., Altmann, J. (2013). Adapting smartphone as learning technology in a Korean university. *Journal of Integrated Design and Process Science*, 17, 5-16. DOI: 10.3233/jid-2013-0002

Kustyorini, Y. (2012). *Pengaruh Pembelajaran Berbasis Masalah Dilengkapi Media Virtual terhadap Aktivitas dan Hasil Belajar Fisika SMA/MA*. S2 Tesis, UNY. Retrieved form https://core.ac.uk/download/pdf/11065836.pdf

Lestari, Fitria & SARYANTONO, Buang & Syazali, Muhamad & Saregar, Antomi & MADIYO, Madiyo & JAUHARIYAH, Durrul & Umam, Rofiqul. (2019). Cooperative Learning Application with the Method of "Network Tree Concept Map":Based on *Japanese Learning System Approach, 7,* 15-32. 10.17478/jegys.471466.

Loucks, S. E. (2007). *Introductory physics with algebra: Mastering problem-solving.* US: John Wiley & Sons.

Mardapi D. (2012). *Pengukuran Penilaian & Evaluasi Pendidikan*. Yogyakarta: Nuha Litera.

Menteri Pendidikan dan Kebudayaan. (2016). Permendikbud Nomor 22 Tahun 2016, tentang Standar Proses Pendidikan Dasar dan Menengah.

Neubert, J., Mainert, J., Kretzschmar, A., & Greiff, S. (2015). The Assessment of 21st Century Skills in Industrial and Organizational Psychology: Complex and Collaborative Problem Solving. *Industrial and Organizational Psychology, 8*(2), 238-268. doi:10.1017/iop.2015.14

Obenchain, K., & Ives, B. (2006). Experiential Education in the Classroom and Academic Outcomes: For Those Who Want it All. *Journal of Experiential Education, 29*(1), 61–77. https://doi.org/10.1177/105382590602900106

Polya, G. (1973). *How To Solve it.* New Jersey: Princeton University Press.

Ruseffendi. (1991). Pengantar kepada Membantu Guru Mengembangkan Kompetensinya dalam Pengajaran Matematika untuk Meningkatkan CBSA. Bandung, Tarsito.

Shafa. (2014). Karakteristik Proses Pembelajaran Kurikulum 2013. Artikel: Dinamika Ilmu, 14(1), 81-95.

Subali, B., & Suyata, P. (2012). *Pengembangan Item Tes Konvergen dan Divergen: Penyelidikan Validitasnya secara Empiris*. Yogyakarta: Diandra Pustaka Indonesia.

Supahar, Prasetyo Z. K.. (2015). Pengembangan instrumen penilaian kinerja kemampuan inkuiri peserta didik pada mata pelajaran fisika SMA. *Jurnal Penelitian dan Evaluasi Pendidikan, 19*(1), 96-108. https://doi.org/10.21831/pep.v19i1.4560

Suryabrata, S. (2002). *Pengembangan Alat Ukur Psikologis.* Yogyakarta: Andi Offset.