AL IBTIDA: JURNAL PENDIDIKAN GURU MI (2021) Vol 8 (1) : 130-143
DOI: http://dx.doi.org/ 10.24235/al.ibtida.snj.v8i1.7865

Al Ibtida: Jurnal Pendidikan Guru MI
ISSN: 2442-5133, e-ISSN: 2527-7227
Journal homepage: http://syekhnurjati.ac.id/jurnal/index.php/ibtida
Journal email: alibtida@syekhnurjati.ac.id

Al Ibtida

Disclosure of Student Ability in Working on Higher-Order Thinking Skills Questions through Rasch Modeling

Ade Yulianto*
*Primary Education Study Program, School of Postgraduate Studies,
Universitas Pendidikan Indonesia, Bandung, Indonesia
Email: adeyulianto@upi.edu

Aan Yuliyanto**
**Primary Education Study Program, School of Postgraduate Studies,
Universitas Pendidikan Indonesia, Bandung, Indonesia
Email: aanyuliyanto@upi.edu

Ghullam Hamdu***
***Primary Teacher Education Study Program,
Universitas Pendidikan Indonesia Tasikmalaya Campus, Tasikmalaya, Indonesia
Email: ghullamh2012@upi.edu

Lutfi Nur****
****Primary Teacher Education Study Program,
Universitas Pendidikan Indonesia Tasikmalaya Campus, Tasikmalaya, Indonesia
Email: lutfinur@upi.edu

Desi Fitriani*****
*****An-Nahl Islamic Elementary School, Tasikmalaya, Indonesia
Email: desifitiriani5757@gmail.com

Noorhisham Hamzah******
******Kebangsaan Beranang School, Hulu Langat Selangor, Malaysia
Email: noorhisham@msil.gurukreatif.net

Received: February 14th, 2021. Accepted: June 12th, 2021. Published: June 27th, 2021.

Abstract

Information on students' ability in higher-order thinking obtained during the learning process can certainly provide an overview for teachers to evaluate appropriate and effective learning. In this context, how can a teacher properly carry out diagnostic and remedial teaching based on the information on student abilities obtained? Therefore, this study aims to provide a disclosure analysis technique for the acquisition of students' abilities in higher-order thinking. This research is one of the stages of Design-Based Research (DBR), which is a reflection to produce design principles and perfect their implementation. Determination of the research sample is done by using the purposive sampling technique. Data collection of students' ability in higher-order thinking was done through HOTS-based test questions developed based on the cognitive hierarchy adopted from Bloom's taxonomy. The results of the acquisition of students' abilities in higher-order thinking were then analyzed through Rasch modeling with the help of the Winsteps 3.75 application. Based on the results of

Rasch modeling, it was obtained that the students' abilities were grouped into high, medium, and low categories and had a level of suitability of abilities that did not need to be reviewed, and there were no biased student abilities. The results of data processing and analysis of students' ability in higher-order thinking have implications for teacher actions in carrying out appropriate and effective learning evaluations as well as mapping students' abilities in higher-order thinking with unbiased conformity.

Keywords: *student ability, HOTS-based test questions, rasch modeling.*

Abstrak

Informasi kemampuan siswa dalam berpikir tingkat tinggi yang diperoleh selama proses pembelajaran tentunya dapat memberikan gambaran bagi guru untuk melakukan evaluasi pembelajaran yang tepat dan efektif. Dalam konteks tersebut, bagaimana seorang guru dapat secara tepat melakukan diagnostic and remedial teaching berdasarkan informasi kemampuan siswa yang diperoleh?. Maka dari itu, penelitian ini bertujuan untuk memberikan teknik analisis pengungkapan pemerolehan kemampuan siswa dalam berpikir tingkat tinggi. Penelitian ini merupakan salah satu tahapan dari *Desain Based Reaserch (DBR)*, yakni refleksi untuk menghasilkan prinsip-prinsip desain dan menyempurnakan implementasinya. Penentuan sampel penelitian dilakukan dengan menggunakan teknik *purposive sampling*. Pengumpulan data kemampuan siswa dalam berpikir tingkat tinggi dilakukan melalui pengerjaan soal tes berbasis HOTS yang dikembangkan berdasarkan hirarki kognitif yang diadopsi dari taksonomi Bloom. Hasil pemerolehan kemampuan siswa dalam berpikir tingkat tinggi kemudian dianalisis melalui pemodelan rasch dengan berbantuan aplikasi winsteps 3.75. Berdasarkan hasil pemodelan Rasch diperolehlah pengelompokkan kemampuan siswa dengan kategori tinggi, sedang, dan rendah serta memiliki tingkat kesesuaian kemampuan yang tidak perlu ditinjau ulang dan tidak terdapat kemampuan siswa yang bias. Hasil pengolahan dan analisis data kemampuan siswa dalam berpikir tingkat tinggi berimplikasi pada tindakan guru dalam melaksanakan evaluasi pembelajaran yang tepat dan efektif serta pemetaan kemampuan siswa dalam berpikir tingkat tinggi dengan kesesuaian yang tidak bias.

Kata kunci: *kemampuan siswa, soal tes berbasais HOTS, pemodelan rasch.*

INTRODUCTION

Higher-order thinking skills are thinking skills that do not only stop at the stage of receiving and storing the information obtained but they also empower ones to use that information to understand oneself about a phenomenon so that it can give birth to action. Higher-order thinking requires a person to apply new information or prior knowledge and manipulate information to reach possible answers in new situations (Heong et al., 2011). In addition, higher-order thinking skills are thinking at a higher level than just memorizing facts or saying something to someone exactly like something that was told to us (Rofiah et al., 2013).

Based on the explanation above, it can be concluded that the ability to think at a high level is not only limited to receiving and storing the information obtained but also being able to empower the information to understand oneself about a phenomenon so that it can give birth to action. The term higher-order thinking skills (HOTS) is part of the classification of thinking skills presented by Bloom. Bloom classifies thinking skills into six levels, namely remembering (C-1), understanding (C-2), applying (C-3), analyzing (C-4), evaluating (C-5)) and creating (C-6) (Bloom et al., 1984). Higher-order thinking which are considered the top end of Bloom's taxonomy are analyzing, evaluating and creating; all of which require complex thinking to solve problems through critical and creative thinking skills (Bestiyana, 2018; Brookhart, 2010; Plan, 2014).

Studies that focus on children say it is important to start developing higher-order thinking skills from an early age. Even children can use many types of higher-order thinking skills at very high levels when given the opportunity, support, guidance, and practice. Developing students' higher-order thinking is a powerful strategy to improve educational outcomes for all students (Saifer, 2018). Higher-order thinking skills can help promote growth and promote academic achievement because students need to think beyond literal questions, so higher-order thinking skills can train students to think on their own (Singh et al., 2020).

Thus, students' HOTS should be measured to be understood and developed from an early age. In measuring HOTS ability, an assessment can be carried out by giving HOTS-based test questions. This assessment means providing good information for teachers to help students learn better. The analysis process of students' HOTS ability acquisition can be done through Rasch modeling assisted by the Winsteps 3.75 application. The Rasch model that provides psychometric analysis techniques can be used by teachers to develop test items and important tools that can provide relevant information related to student assessments for learning (Sumintono, 2018). Analysis of this test instrument using the Rasch model is included in the response item measurement theory. This measure describes the interaction between the subject and the test item. This will make the measurement have more precise and objective result (Sumintono & Widhiarso, 2014). Additionally, the Rasch model is a well-studied measurement approach that models the relationship between item difficulty, people's abilities, and the probability of a given response (Andrich, 1981). The advantage of the Rasch model compared to classical theory is that it can identify incorrect answers from experts, identify incorrect judgments, and predict missing data based on systematic response patterns (Fahmina et al., 2019; Goodwin & Leech, 2003; Ratna et al., 2017).

Research that has been done by previous researchers related to HOTS analysis and Rasch modeling, including the HOTS assessment of high school students in economics lessons through Multidimensional Item Response Theory (MIRT), shows that MIRT offers accurate measurements in estimating multidimensional test parameters. The items have a discriminant multidimensional average and a moderate difficulty level , while students have moderate HOTS abilities. Their ability to think creatively is lower than the ability to think critically and solve problems. The test proved reliable with an Alpha coefficient of 0.81, resulting in the highest information function of 4.0124 and a low measurement error of 0.4992. Therefore, it is suitable to be tested on students with moderate problem solving and critical thinking skills but with high creative thinking skills (Friyatmi, 2020). Based on this research, the MIRT system is used to measure and describe the HOTS of high school students in economics learning.

In addition, Rasch Modeling is implemented to describe the Self-Directed Learning (SDL) of prospective teachers in Papua while learning through the Moodle e-learning platform and its measurement instruments. The results of the reliability and validity analysis showed that the SDL instrument was appropriate. Furthermore, the results of the univariate analysis of the respondent's level of agreement showed a logit value > 0.00, which means that students generally agree with the statement of the SDL instrument obtained during the use of the HOTS-oriented Moodle e-learning platform. In addition, the statement items have a low level of acceptance which is generally related to the ability to measure the level of failure in learning. These results can be used as a reference for lecturers to improve the optimization of e-learning, one of which is through the presentation of appropriate evaluation instruments. So that students, especially prospective teachers, have the ability to measure their learning achievement (Yusuf et al., 2021).

Rasch modeling in subsequent studies has been used to test the motor ability assessment system in early childhood. This study demonstrates the homogeneity of items and people invalidated motor scores, using 12 items from MOT 4-6. Thus, it provides evidence of a single latent construct (i.e., motor competence), which underlies motor skill performance in early childhood. This indicates that calculating the number of successful passes may not be suitable as a scoring system in motor competency assessment. The current findings also support the use of validated composite scores in motor assessment (Utesch et al., 2016). Thus, the study that has been carried out shows that Rasch modeling is used to validate an instrument for measuring SDL and an assessment system in learning for students and early childhood.

It is related to Rasch modeling, which in another study was used to understand students' strategies and difficulties with graph interpretation in three different domains: mathematics, physics (kinematics), and contexts other than physics. Analysis of the difficulty of the questions obtained through Rasch modeling shows that the difficulty level of questions involving context (both physics and other contexts) is higher than math problems directly on the graph. In addition, students' explanations were analyzed and categorized. Students' strategies of graph interpretation were found to be mostly domain-specific. In physics, the dominant strategy seems to be the use of formulas, especially among students at the University of Zagreb. This strategy seems to hinder the more productive use of other strategies, which students in other domains use. Students are generally better at interpreting the slope of the graph than the area under the graph, which is difficult for students and requires more attention in teaching physics and mathematics (Ivanjek et al., 2017). The study showed that Rasch modeling was used to analyze the difficulties and strategies of students in solving problems.

Meanwhile, for elementary school students, Rasch modeling can be used to identify students' academic potential, such as the findings of 19 Aboriginal Canadian children in grades 3 and 4 who were taken through the Coolabah Dynamic Assessment (test-intervention-retest) process in this pilot study, eight made a profit significantly from pre-test to post-test. Among the "underachievers" group, three demonstrated exceptional potential with post-test raw scores indicating high academic potential. In the context of this study, "unseen low achievement" refers to individuals who perform poorly both in class and to commonly used evidence of potential for higher attainment. The profiles of these three students illustrate the value of dynamic assessments in identifying the talents of low-achieving students, including those from disadvantaged and minority group backgrounds (Chaffey et al., 2006). Thus, Rasch modeling applied to elementary school students seems able to describe the criteria for students' abilities. Different from several studies that have been developed, this research focuses on revealing how students' abilities in problems that contain higher-order thinking skills in elementary school are through Rasch modeling.

METHODS

The research method used is Design-Based Research (DBR). DBR can be defined as “a series of approaches, with the intent of producing new theories, and potentially impact learning and teaching in naturalistic settings with the aim of improving educational practice” (Abdallah, 2014; Herrington et al., 2007). More specifically, this article describes one of the stages of the DBR method, namely, Reflection to produce “design principles” and enhance solution implementation. This process describes the results of the implementation of higher-order thinking skills (HOTS) based test questions on the benefits of alternative energy in elementary schools. The determination of the research sample was carried out using the purposive sampling technique where the sampling of data sources was based on certain considerations. This study was conducted on 20 grade 4 elementary school students in the city of Tasikmalaya, West Java Province, Indonesia. Furthermore, the data in this study were collected using a test technique, where students worked on 11 HOTS-based test questions. The HOTS-based test questions were developed by referring to the following grid of questions:

Table 1. HOTS-based Question Grid

Number of Question	Indicator	Cognitive Level
1	Separating the materials used by parachutes by type of natural resource.	C4, Analysis
2	Select the correct statement about the meaning of energy.	C5, Evaluation
3	Reviewing the use of alternative wind energy.	C4, Analysis
4	Reviewing how nature-based parachutes work.	C4, Analysis
5	Making a parachute according to the instructions text.	C6, Create
6	Evaluating the sequence of 5 steps of making a parachute wing based on the drawing.	C5, Evaluation
7	Separating activities that utilize solar thermal energy.	C4, Analysis
8	Reviewing the workings of the solar oven.	C4, Analysis
9	Selecting the use of materials in the solar oven.	C6, Create
10	Separating the right statements about being environmentally friendly.	C4, Analysis
11	Distinguishing characteristics of solar ovens.	C4, Analysis

After developing the HOTS-based test questions, then the instrument was tested by measuring its validity and reliability. The data obtained from the trial results were then processed using the SPSS version 25 application to calculate the validity of the questions using the person correlation value and calculate the reliability of the questions using the Cronbach's Alpha value.

Table 2. HOTS-Based Test Question Validation Results

Number of Questions	**Person Correlation (r_{hitung})**	**r_{tabel} (α=5%)**	**Interpretation**
1	0,836		Valid
2	0,543		Valid
3	0,528		Valid
4	0,631		Valid
5	0,633		Valid
6	0,478	0,444	Valid
7	0,657		Valid
8	0,489		Valid
9	0,684		Valid
10	0,542		Valid
11	0,787		Valid

Tabel 3. Reliability Results of HOTS-Based Test Questions

Reliability Statistics

Cronbach's Alpha	N of Items
.741	11

Table 2 informs that the 11 test items that have been tested are declared valid with a calculated r-value > r-table on all items. Meanwhile, in Table 3, Cronbach's Alpha value has a value of 0.741, which means that the test questions are declared to have acceptable reliability (Gliem & Gliem, 2003).

Furthermore, the process of disclosing students' abilities in working on HOTS-based test questions was carried out using RASCH modeling assisted by Winsteps 3.75 application. The stages of the analysis process are through Rasch modeling with the adaptation of stages carried out by (Hamdu et al., 2020), namely:

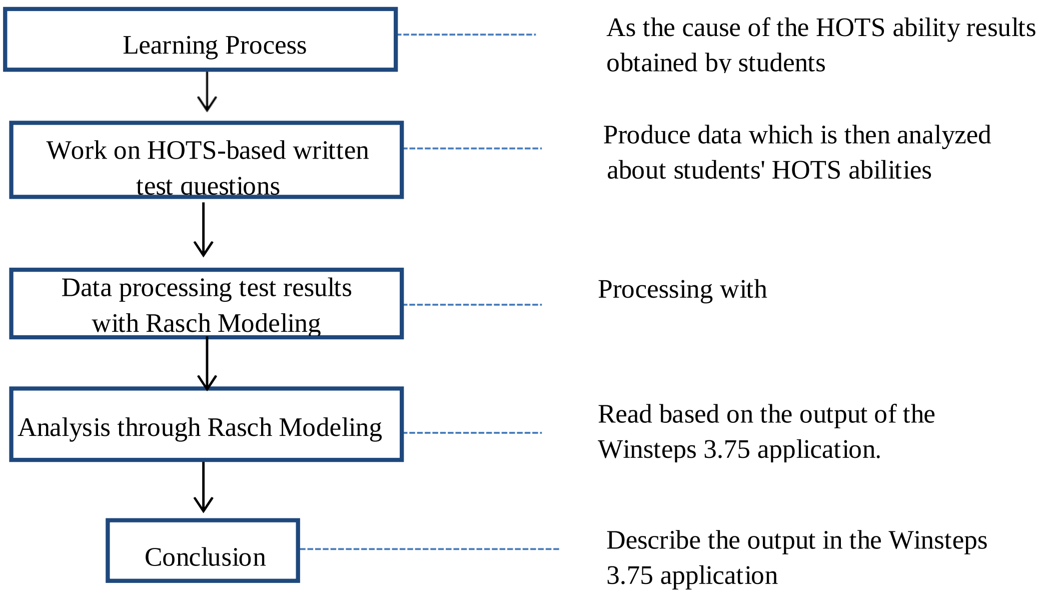


Figure 1. Data Collection & Analysis Process

The stages of the data collection and analysis process are presented in Figure 1. Data collection begins with the implementation of learning activities as the cause of the results of acquiring HOTS abilities in students. After learning is complete, students are directed to work on HOTS-based test questions that produce data on students' HOTS abilities. Then the data is processed through Rasch modeling with Winsteps 3.75, which provides an overview to describe and analyze students' ability level and suitability level.

RESULTS AND DISCUSSION

The study focuses on the ability of elementary school students to work on HOTS questions on alternative energy utilization materials andthe results of students' abilities are revealed through an analysis process using Rasch modeling assisted by the Winsteps 3.75 application. Rasch modeling in the process of disclosing students' abilities has been widely used by previous researchers, among others, to measure students' critical thinking skills in STEM learning in elementary schools (Hamdu et al., 2020). In addition, it is used to develop mathematical anxiety instruments, character, and student confidence in elementary school students (Karlimah et al., 2020; Nur et al., 2020; Rusmana et al., 2020). Analysis of students' abilities in working on HOTS-based test questions can identify students' ability levels in working on these questions. Data on the level of student ability can be seen through the analysis in Table 4.

Table 4. Analysis of Student Ability Levels

Total Score	Total Count	Measure	Person
11	11	4.23	P14
11	11	4.23	P16
11	11	4.23	P17
10	11	2.81	P10
10	11	2.81	P15
8	11	1.17	P01
8	11	1.17	P02
7	11	0.64	P04
7	11	0.64	P06
6	11	0.17	P13
5	11	-0.28	P03
5	11	-0.28	P09
4	11	-0.73	P12
4	11	-0.73	P19
3	11	-1.22	P05
3	11	-1.22	P11
3	11	-1.22	P18
2	11	-1.82	P07
2	11	-1.82	P20
1	11	-2.69	P08
Mean		0.50	
S.D.		2.90	

Information on the grouping of students' ability levels in doing HOTS-based test questions can be seen through the standard deviation (SD) value and the starting point of the average logit person value (Sumintono & Widhiarso, 2015). Based on the results of the Rasch model through the use of Winsteps 3.75 software, the students' abilities were grouped into: very high, high, medium, and low. This grouping is based on the SD score (standard deviation = 2.09) and the MEAN value (0.50) as a reference for identifying student groupings. Then the grouping value range is obtained as follows: if the student's ability> SD (2.09), then the student has the high ability, if SD (2.09) <student ability <MEAN (0.50), then the student has the moderate ability, If the student's ability <MEAN (0.50), then students have the low ability.

From the results of the range provisions above, the student's ability can be described as follows:

1) The high ability level category consists of 5 students (P14, P16, P17, P10 & P05) based on their measure (logit value).
2) The moderate ability level category consists of 4 students (P01, P02, P04, & P06) based on their measure (logit value).
3) The low ability level category consists of 11 students (P13, P03, P09, P12, P19, P05, P11, P18, P07, P20, & P08) based on their measure (logit value).

After categorize students 'abilities into high, medium, and low groups, an analysis of the level of students' abilities suitability was carried out by detecting student response patterns in working on HOTS-based test questions. The results of this analysis enable us to see patterns of inappropriate responses by students in answering questions based on students' abilities that had been previously analyzed. Figure 3 shows the information regarding the suitability level of students' abilities in doing HOTS-based test questions.

The criteria used to check the suitability of students' abilities are the same as the criteria for checking the suitability of the items (outliers or misfits). (Boone et al., 2013), namely:

1) Accepted outfit means-square (Outfit MNSQ) value: 0.5 <MNSQ <1.5.
2) Accepted Z-Standard Outfit Value (ZSTD Outfit): -2.0 <ZSTD <+2.0.
3) Point Measure Correlation (PT-Measure Corr) value: 0.4 <Pt-Measure Corr <0.85.

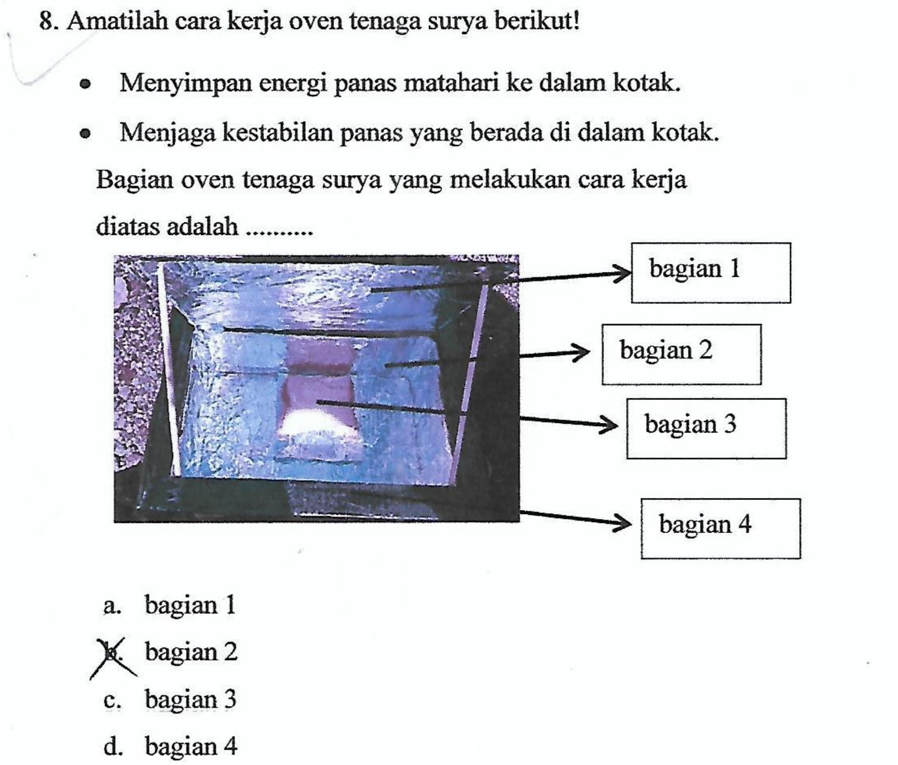
Table 5. Analysis of Ability Student fit order level

Outfit		**Pt-Measure**	**Person**
MNSQ	**ZSTD**	**Corr.**	
9.90	4.0	-0.75	P08
1.81	1.0	-0.12	P20
1.55	0.9	-0.09	P18
1.27	0.7	0.08	P10
1.26	0.7	0.17	P13
1.19	0.5	0.33	P01
0.97	0.1	0.33	P19
ghi0.78	-0.1	0.41	P11
0.85	-0.2	0.50	P09
0.70	-0.4	0.58	P02
0.57	-0.2	0.46	P07
0.62	-0.6	0.62	P12
0.62	-0.9	0.62	P03
0.57	-1.0	0.57	P06
0.48	-0.6	0.48	P05
0.50	-1.2	0.50	P04
0.16	-0.5	0.16	P15

Based on the data in Table 5, the MNSQ scores of students P08, P20 & P18 were not accepted; the ZSTD score of student P08 was not accepted; all students' Pt-Measure Corr scores were accepted. If the student's ability category on the three criteria (MNSQ, ZSTD, and Pt. Measure Corr) is not met, certainly it is not suitable and needs to be reviewed or there is a biased ability (Bond & Fox, 2013; Boone et al., 2013). From these quotations, all students are within the limits of the reasonableness of their response patterns, or in other words, they have an acceptable level of conformity. The results of the student's ability suitability analysis are of course based on the statistical description of Rasch's suitability which can provide a useful framework for testing the correctness of a person's response, measuring the estimation of a person's response ability and being able to detect various disturbances to a person's response (Smith, 1986). In addition, the results of the analysis show that the Rasch model is written as a probability model of individual responses to an item and therefore does not explicitly become a response model itself (Brogden, 1977).

Based on the analysis results, it was found that Rasch modeling was able to identify students' abilities and categorize them into high, medium, or low abilities. The results of this ability category are devoted to the results of the HOTS-based written test questions. These results indicate that the student's ability to work on HOST-based questions is more at low ability. The learning that has been carried out is only able to produce 5 students out of 20 students (25%) with high ability level. Meanwhile, in the medium and low category, there are

4 students (20%) and 11 students (55%) respectively . Therefore, these results indicate that it is necessary to optimize students' acquisition of HOTS abilities, especially as a reflective material for teachers in optimizing the implementation of learning. Learning is done by stimulating students to have access to a variety of sources relevant to the issues examined, which facilitate understanding and ensure the completion of the task (Gorghiu et al., 2015). A task that will be developed by the teacher with the orientation to improve students' thinking processes (Carless, 2004; Kang, 2017).


8. Amatilah cara kerja oven tenaga surya berikut!

- Menyimpan energi panas matahari ke dalam kotak.
- Menjaga kestabilan panas yang berada di dalam kotak.

Bagian oven tenaga surya yang melakukan cara kerja diatas adalah ....

bagian 1
bagian 2
bagian 3
bagian 4

a. bagian 1
b. bagian 2
c. bagian 3
d. bagian 4

Figure 2. Student Test Result

Figure 2 shows the results of answering HOTS-based test questions by student P14 in reviewing how solar power works. The question describes the part of the solar oven that stores solar thermal energy and maintains the heat stability in it. Students P14 was able to correctly analyze how the solar oven works in storing solar thermal energy and maintaining heat stability in it by answering option "b. part 2". Based on the analysis through Rasch modeling, student P14 had high abilities in the HOTS-based test and demonstrated an authentic level of ability adjustment. The results of the analysis on students in answering questions certainly show that Rasch modeling formulates a model that relates students and test items (Sumintono & Widhiarso, 2014).

CONCLUSION

Technical disclosure of students' abilities through Rasch modeling can provide an alternative way of data processing. The results of the analysis of students' abilities in working on HOTS-based test questions provide information that can be used as reflection materials for teachers in the process of implementing learning. The results of data processing and analysis of students' ability in higher-order thinking have implications for teacher actions in carrying out appropriate and effective learning evaluations as well as mapping students' abilities in higher-order thinking with unbiased conformity.

REFERENCES

Abdallah, M. (2014). Design-Based Research (DBR) in Educational Enquiry and Technological Studies: A Version for PhD Students Targeting the Integration of New Technologies and Literacies into Educational Contexts. *Online Submission*.

Andrich, D. (1981). Book Review: Probabilistic Models for Some Intelligence and Attainment Tests (expanded edition: Georg Rasch Chicago: The University of Chicago Press, 1980, 199 pp., 15hardcover, 7 paperback. *Applied Psychological Measurement*, *5*(4), 545–550.

Bestiyana, R. A. (2018). Profil Berpikir Kritis Siswa Smp Dalam Menyelesaikan Soal Higher Order Thinking Matematik Ditinjau Dari Gaya Kognitif Visualizer-Verbalizer. *MATHEdunesa*, *7*(1).

Bloom, B. S., Krathwohl, D. R., & Masia, B. B. (1984). Bloom taxonomy of educational objectives. In *Allyn and Bacon*. Pearson Education.

Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model: Fundamental measurement in the human sciences*. Psychology Press.

Boone, W. J., Staver, J. R., & Yale, M. S. (2013). *Rasch analysis in the human sciences*. Springer.

Brogden, H. E. (1977). The Rasch model, the law of comparative judgment and additive conjoint measurement. *Psychometrika*, *42*(4), 631–634.

Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. ASCD.

Carless, D. (2004). Issues in teachers' reinterpretation of a task-based innovation in primary schools. *Tesol Quarterly*, *38*(4), 639–662.

Chaffey, G. W., Halliwell, G., & McCluskey, K. W. (2006). Identifying high academic potential in Canadian Aboriginal primary school children. *Gifted and Talented International*, *21*(2), 61–70.

Fahmina, S. S., Masykuri, M., Ramadhani, D. G., & Yamtinah, S. (2019). Content validity uses Rasch model on computerized testlet instrument to measure chemical literacy capabilities. *AIP Conference Proceedings*, *2194*(1), 20023.

Friyatmi, M. (2020). Assessing Students' Higher Order Thinking Skills Using Multidimensional Item Response Theory. *Problems of Education in the 21st Century*, *78*(2), 196.

Gliem, J. A., & Gliem, R. R. (2003). *Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales*.

Goodwin, L. D., & Leech, N. L. (2003). The meaning of validity in the new standards for educational and psychological testing: Implications for measurement courses. *Measurement and Evaluation in Counseling and Development*, *36*(3), 181–191.

Gorghiu, G., Drăghicescu, L. M., Cristea, S., Petrescu, A.-M., & Gorghiu, L. M. (2015). Problem-based learning-an efficient learning strategy in the science lessons context. *Procedia-Social and Behavioral Sciences*, *191*, 1865–1870.

Hamdu, G., Fuadi, F. N., Yulianto, A., & Akhirani, Y. S. (2020). Items Quality Analysis Using Rasch Model To Measure Elementary School Students' Critical Thinking Skill On Stem Learning. *JPI (Jurnal Pendidikan Indonesia)*, *9*(1), 61–74.

Heong, Y. M., Othman, W. B., Yunos, J. B. M., Kiong, T. T., Hassan, R. Bin, & Mohamad, M. M. B. (2011). The level of marzano higher order thinking skills among technical education students. *International Journal of Social Science and Humanity*, *1*(2), 121.

Herrington, J., McKenney, S., Reeves, T., & Oliver, R. (2007). Design-based research and doctoral students: Guidelines for preparing a dissertation proposal. *EdMedia+ Innovate Learning*, 4089–4097.

Ivanjek, L., Planinic, M., Hopf, M., & Susac, A. (2017). Student difficulties with graphs in different contexts. In *Cognitive and affective aspects in science education research* (pp. 167–178). Springer.

Kang, H. (2017). Preservice teachers' learning to plan intellectually challenging tasks. *Journal of Teacher Education*, *68*(1), 55–68.

Karlimah, K., Andriani, D., & Suryana, D. (2020). Development of Mathematical Anxiety Instruments with a Rasch Model Analysis. *The Open Psychology Journal*, *13*(1).

Nur, L., Nurani, L. A., Suryana, D., & Ahmad, A. (2020). Rasch model application on character development instrument for elementary school students. *International Journal of Learning, Teaching and Educational Research*, *19*(3), 437–459.

Plan, Q. E. (2014). *Improving students' higher-order thinking competencies, including critical evaluation, creative thinking, and reflection on their own thinking*. January.

Ratna, I., Yamtinah, S., Ashadi, M., Masykuri, M., & Shidiq, A. (2017). The Implementation of Testlet Assessment Instrument in Solubility and Solubility Product Material for Measuring Students' Generic Science Skills. *International Conference on Teacher Training and Education 2017 (ICTTE 2017)*.

Rofiah, E., Aminah, N. S., & Ekawati, E. Y. (2013). Penyusunan Instrumen tes kemampuan berpikir tingkat tinggi fisika pada siswa SMP. *Jurnal Pendidikan Fisika*, *1*(2).

Rusmana, N., Hafina, A., Wardhany, R. O., & Suryana, D. (2020). Students' Confidence Instrument Analysis in Poetry Learning through Rasch Model. *The Open Psychology Journal*, *13*(1).

Saifer, S. (2018). *HOT skills: Developing higher-order thinking in young learners*. Redleaf Press.

Singh, C. K. S., Gopal, R., Ong, E. T., Singh, T. S. M., Mostafa, N. A., & Singh, R. K. A. (2020). Teachers Strategies to Foster Higher-Order Thinking Skills To Teach Writing. *Malaysian Journal of Learning and Instruction*, *17*(2), 195–226.

Smith, R. M. (1986). Person fit in the Rasch model. *Educational and Psychological Measurement*, *46*(2), 359–372. https://doi.org/10.1177%2F001316448604600210.

Sumintono, B. (2018). Rasch Model Measurements as Tools in Assesment for Learning. *1st International Conference on Education Innovation (ICEI 2017)*, 38–42.

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

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan rasch pada assessment pendidikan*. Trim komunikata.

Utesch, T., Bardid, F., Huyben, F., Strauss, B., Tietjens, M., De Martelaer, K., Seghers, J., & Lenoir, M. (2016). Using Rasch modeling to investigate the construct of motor competence in early childhood. *Psychology of Sport and Exercise*, *24*, 179–187.

Yusuf, I., Widyaningsih, S. W., Prasetyo, Z. K., & Istiyono, E. (2021). The analysis of self directed learning (SDL) through Rasch modeling: Case study on prospective teachers during the use of e-learning with HOTS-oriented in the period of Covid-19 pandemic. *AIP Conference Proceedings*, *2330*(1), 50006.