A Rasch Model Analysis on Junior High School Students' Scientific Reasoning Ability

Mustika Wati, Saiyidah Mahtari ([✉]), Sri Hartini, Heny Amalia
FKIP Universitas Lambung Mangkurat, Indonesia
`saiyidah_pfis@ulm.ac.id`

Abstract—Science education that emphasizes only the count is not relevant at the moment. Students must make scientific reasoning in answering the problem. This study aims to determine students' scientific reasoning abilities in light matter. The method in this research was the descriptive method by using the survey technique. The instrument used is a description test consisting of 8 items of light matter. This test was conducted on 201 students from the eighth-grade junior high school in Banjarmasin selected at random. The RASCH model is used as a processing stage of data from raw data into logit numbers that provide information related to infit, outfit and unidimensionality using a winstep program to achieve this goal, this study investigated the quality of items from Item and person measure, Item Bias, item and person reliability, and variable map. The findings show that the scientific reasoning ability of the eighth-grade students can be at a low level. So researchers should strive to improve students' scientific reasoning abilities in future research.

Keywords—Rasch Model, Construct Validity, Test, Scientific Reasoning

1 Introduction

Physics can be interpreted as a science of measurement because everything we know about the world of physics and the principles governing its behaviour has been studied through observations of natural phenomena. Physics is a science that requires more understanding than shipping (Siregar, 2003). The purpose of learning physics is the formation of reasoning ability in students reflected through the ability to think logically, critically, and systematically in problems solving, especially in the field of physics (Rangkuti, 2015). In the Curriculum 2013, it is explained that one of the core competencies in learning, especially for grades VIII and IX SMP is to cultivate, decorate, and reason in the realm of concrete and abstract realms as studied in schools and other sources in the same point of view/theory.

Scientific reasoning is one of the 21st-century skills that is expected to be taught in the science classroom in an effort to prepare students for their success in facing the challenges of globalization. Scientific reasoning is highly emphasized in new science education standards (Zhou, et al., 2016). In the PISA test, the skill is also one of the skills tested (Salz, 2009). The scientific reasoning could be an effective predictor of

student success and thus could potentially be used in practical decision making for the course (Thompson, et al., 2018). Students who are used to solve problems indirectly develop the thought process of reasoning (Rizta et al., 2013). In a series of studies, they showed that the scientific reasoning of preadolescent children was severely deficient (Kuhn & Franklin, 2006).

Georg Rasch developed an analytical model of the response theory of grains (or Item Response Theory, IRT) in the 1960s (Boone et al., 2011). IRT is an alternative test measurement theory in addition to Classical test theory. Classical tests of these theories enable the presumption of test results such as the difficulty of items and the ability of people. IRT focuses on the pattern of responses given by the person to the test item and background person. IRT has many advantages and is more complex than the Classical test theory (Chan et al., 2013). Rasch measurement model or one parameter model is the simplest IRT model, and it has strong measurement properties (Afrassa, 2005). Raw data in the form of dichotomous data (in the form of right and wrong) that indicate students' abilities, Rasch formulates this into a model that connects students and item (Sumintono and Widhiarso, 2014). The benefit of Rasch's analysis is its ability to estimate the total score for clients even when not all items have been managed. This is especially useful when working with children who may not adhere to all the constraints of standard testing situations (Avery et al., 2003).

Scientific reasoning is very important for students. It is related to the role of scientific reasoning in the process of solving problems in physics. Thus, this study was conducted to describe the students' scientific reasoning abilities using the Rasch measurement model.

2 Methodology

The method in this research was the descriptive method by using the survey technique. In this study, the participant was 201 eighth grade students from 3 junior high schools in Banjarmasin. They were chosen at random. Their age was between 14-15 years old. The study was conducted in April 2017. All participants had studied light material before the test. The instrument used in this study is a scientific reasoning test on light material with eight questions. The purpose of this test was to test students' scientific reasoning abilities. Students are given time to answer the test for eighty minutes. The test results are then used as data in this study. Each student is labelled with codes A, B and C for the representation of the origin of the school. The description described students' reasoning abilities gained after students completed the reasoning ability of reasoning instruments on the Light material. The result data will be included in the winstep software which is one of the series in the Rasch model with the data polytomy. Output in this software that is in the form of the Item Measure table, Person Measure, Variable Maps and Reliability which have been converted before becoming a logit number. This logit number must qualify the Mean Square Output (MNSQ), Outfit Z-Standard (ZSTD), Point Measure Correlation (Pt Mean Corr) and Reliability values according to Rasch modelling.

Output winstep is in the form of Item Measure table, Person Measure, Variable Maps and Reliability which have been converted before becoming logit number. This logit number must qualify the Mean Square Output (MNSQ), Outfit Z-Standard (ZSTD), Point Measure Correlation (Pt Mean Corr) and Reliability values according to Rasch modelling. The parameters used are infit and outfit of the mean square and standardized values. According to Sumintono and Widhiarso (2014), infit (inlier sensitive or information weighted fit) is the sensitivity of response pattern to target item on respondent (person) or vice versa; while outfit (outlier sensitive fit) measures the sensitivity of the response pattern to the item with a certain degree of difficulty on the respondent or vice versa. According to Sumintono and Widhiarso (2015) of the logit numbers obtained from the Ministep software output, there is an interval scale (logit rule) which describes the state of the number. The scale is:

- The value of Mean Square Outfit (MNSQ) received: 0.5 <MNSQ <1.5
- The accepted Z-Standard Output (ZSTD) value: -2.0 <ZSTD <+2,0
- Point Measure Correlation Value (Pt Mean Corr): 0.4 <Pt Measure Corr <0.85

Therefore, the respondent would be qualified based on the value relating to the entry and absence of respondents in the modelling. Reliability questions with Rasch modelling were analyzed using individual separation values and grain separations as well as Cronbach Alpha values displayed in Rasch program outputs. The higher the value of individual separation and the value of the grain separation and the Alpha Cronbach value the better the reliability of the problem. The criteria for interpreting the value of individual separation and the separation of an instrument's grain can be seen using Table 1 (Sumintono and Widhiarso, 2015).

Table 1. Interpretation of individual separation values and Instrument separation items

Criteria	Value of separation individual and item
Weak	<0,67
Enough	0,67-0,80
Nice	0,81-0,90
Very Good	0,91-0,94
Special	>0,94

The Cronbach Alpha values used to measure interactions between individuals with whole grains can be interpreted using Table 2 (Sumintono and Widhiarso, 2015).

Table 2. Interpretation of Cronbach Alpha values

Criteria	Alpha Cronbach
Bad	< 0,5
Ugly	0,5 – 0,6
Enough	0,6 – 0,7
Nice	0,7 – 0,8
Very Good	>0,8

Problem level of difficulty with Rasch modelling is analyzed by using logit number contained in the measurement column problem, the higher the logit value than the higher the difficulty level of the problem. In the measurement of the problem, there is also information on the standard deviation value. If the value is combined with an average logit score, the difficulty level of the items can be grouped according to the difficulty level as in Table 3 (Sumintono and Widhiarso, 2015).

Table 3. Interpretation of Problem Exchange Index

Interpretation	Criteria
Difficult	0,00 logit + 1SD
Very Difficult	>+ 1SD
Easy	0,00 logit – 1SD
Very Easy	<- 1SD

3   Result and Discussion

Table 4. Results summary of the winstep program output

Item number	Infit		Outfit		PT-Measure		Prob	Measure
	MNSQ	ZSTD	MNSQ	ZSTD	CORR	EXP		
1	1.15	1.5	1.15	0.8	0.59	0.64	0.4915	-0.81
2	1.03	0.3	0.92	-0.6	0.68	0.67	0.0001	-0.26
3	0.79	-2.0	0.81	-1.7	0.69	0.65	0.2022	0.13
4	1.05	0.5	1.22	1.5	0.55	0.6	0.5853	-0.20
5	0.61	-2.8	0.6	-2.7	0.77	0.7	0.2197	0.64
6	0.5	-4.1	0.52	-3.7	0.81	0.71	0.4565	0.63
7	1.38	2.7	1.47	2.7	0.49	0.58	0.0031	-0.14
8	1.18	1.3	1.29	1.6	0.6	0.63	0.0111	0.01
S.D								0.45

Person Reliability 0.79
Iem Reliability 0.99
Cronbach Alpha 0.85

```
                        Person - MAP - Item
                          <more>|<rare>
  1                            +
                           A  |T
                              |
                              |  E5     E6
                           T  |S
                      A A A A  |
                           A  |
                      A A A A  |  E3
  0             A A A A A A A A A S+M E8
          A A A A A A A A A A A B B C C  |  E7
A A A A A A A A A A A A B C C C C C C C C C C C C  |  E2     E4
    A A A A A A A A A B B B B B B C C C  |
    A A A A A B B B B B B B B B B B B B C  |S
    A A A B B B B B B B B B B B B B C C C C M|
        A A A B B B B B B B B B B B B C  |  E1
    A B B B B B B B B B B B B B B B B B B C  |T
 -1                    B B B B B B  +
                  A B B B B C C S|
                      B B B C  |
                          B C  |
                          B B  |
              B B B B B B B C  |
                           T|
                  B B C C C  |
 -2                            +
                              |
                          C  |
                              |
                              |
                              |
 -3                            +
                              |
                      B C  |
                              |
                              |
                              |
 -4                            +
                              |
                              |
                              |
                              |
 -5                    B  +
```
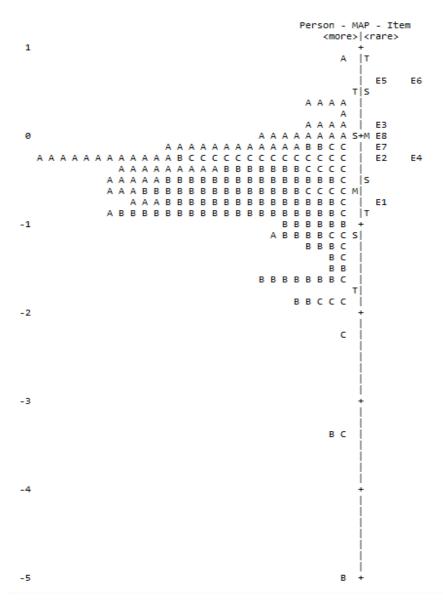
Fig. 1. Variable map

Based on Table 4, it can be seen that the top item is item 7 has a tendency that is less fit. When viewed from three criteria, item 7 only does not qualify at ZSTD outfit value that was equal to 2.7 or unpredictable data, but for MNSQ outfit value that was equal to 1.47 still, meet good criteria for measurement. The value of Pt Measure Corr also meets the criteria of 0.49. Therefore, item 7 can be maintained and did not need to be fixed. This is similar to the items 5 and six that have ZSTD values of -2.7 and -3.7 or too small, so data was too predictable but can still be maintained to measure

students' reasoning abilities. As for the items other items have the value of MNSQ, ZSTD, and Pt Measure Corr by the criteria, so it does not need to be repaired. However, according to Sumintono and Widhiarso (2015), the value of the grain conformity consisting of Outfit MNSQ, ZSTD, and Pt Measure Corr is strongly influenced by the size of the sample size.

A point item is called bias if it is found that one individual with a particular characteristic is more favourable than that of an individual with another characteristic (Sumintono and Widhiarso 2015). An item is said to contain bias if it is found that the probability value of the item is below 5%. In Rasch modelling to detect the biased problem can be seen in Table 4. There were three items that are biased, and five items are unbiased. So, there were five items that can be directly used to measure students' reasoning ability and three items that need to be improved to make good measurements were items 2, 7 and 8. This indicates that these three items need to be improved so as not to harm a particular school.

Reliability problems developed for the criteria of person reliability into the category was nice, the criteria of the reliability of the item entered into the category special, and Cronbach Alpha criteria into the category of very good Thus, overall these questions can be trusted to measure students' reasoning abilities. It also shows that the reliability of the items was very good. In Table 4, it can be seen in the Item Number column that item 5 was the hardest problem with a logit value of 0.64 and question number 1 was the easiest problem with a logit value of -0.81, it corresponds to the cognitive domain of the reasoning problem developed. Items 5, 6, and 3 had Cognitive domains C5, items 8, 7, and 4 had C4 cognitive domains, and item 1 has C3 cognitive domains. Item 2 had Cognitive domain C5, but it was not by the results of research analysis that shows the problem is entered in the category very easily with a logit value of -0.26. It was also shown by the number of students who can correctly answer the question.

The variable map shows the distribution of student ability and item difficulty on the same logit scale. Students' abilities are listed on the left side of the map while the item difficulty is on the right side of the map. The higher logit represents students with higher abilities (the left side) and more difficult items (right side) and vice versa (Iramaneerat, Smith, and Smith, 2008). Through the variable map, it allows us to identify whether the item matches the student's abilities. The relation of ability possessed by students with problem level in figure 1, on the left side shows the distribution of students' reasoning ability and the right side shows the difficulty level of the item. The left side of the map shows the higher level of the reasoning problem. This means the student could get the maximum value of all questions.

On the right side of the map are eight items that have difficulty levels that from item 5 was the most difficult to item1 the easiest to do. This means that there is no problem that accumulates in one line only, it shows the problem of having various levels of difficulty, ranging from the most difficult to the easiest. Logit 0 is set as the average test item (Iramaneerat, Smith, and Smith, 2008). From the variable map, we can see that most students are below the average of the exam items. Few students with higher abilities are above logit 0 and very many students are below average. Low logit indicates low ability. Thus we can argue that the students' ability in scientific reason-

ing is low, as most of them cannot solve the problem of scientific reasoning. In other words, items are deemed less able to fit the student's abilities. This is because students are not familiar with the scientific reasoning items in this study and they are not taught to answer these kinds of questions in school.

Training in scientific reasoning may also have a long-term impact on student academic achievement (Bao, et al., 2009). Scientific reasoning has an important role in the problem-solving process (Khan and Ullah,2010). When students have high problem-solving skills, it can have an impact on achieving more effective student learning outcomes (Nieminem et al., 2012; Stephens and Clemen, 2010). Low scientific reasoning ability makes learning outcomes low, so media, learning and teaching model, and learning materials are needed to improve student learning outcomes. This is supported by the results of the study (Wati, et al., 2018; Erika, et al., 2018; Jatmiko et al., 2018; Limatahu, et al., 2018; Prahani, et al, 2016; Prahani, et al., 2018; Sunarti, et al., 2018; Suyidno et al., 2018) that the media, learning and teaching model that is qualified will be able to improve and achieve the learning outcomes.

4   Conclusions

Scientific reasoning abilities among eighth-grade junior high school students are still at a low level. It is, therefore, necessary to improve students' scientific reasoning.

5   References

[1] Afrassa, T. M. (2005). Monitoring mathematics achievement over time. In *Applied Rasch Measurement: A Book of Exemplars* (pp. 61-77). Springer, Dordrecht. https://doi.org/10.1007/1-4020-3076-2_4

[2] Avery, L. M., Russell, D. J., Raina, P. S., Walter, S. D., & Rosenbaum, P. L. (2003). Rasch analysis of the gross motor function measure: validating the assumptions of the Rasch model to create an interval-level Measure1. *Archives of Physical Medicine and Rehabilitation*, *84*(5), 697-705. https://doi.org/10.1016/s0003-9993(02)04896-7

[3] Bao, L., Cai, T., Koenig, K., Fang, K., Han, J., Wang, J., ... & Wang, Y. (2009). Learning and scientific reasoning. *Science*, *323*(5914), 586-587.

[4] Boone, W. J., Townsend, J. S., & Staver, J. (2011). Using Rasch theory to guide the practice of survey development and survey data analysis in science education and to inform science reform efforts: An exemplar utilizing STEBI self-efficacy data. *Science Education*, *95*(2), 258-280. https://doi.org/10.1002/sce.20413

[5] Chan, S. W., Ismail, Z., & Sumintono, B. (2014). A Rasch model analysis on secondary students' statistical reasoning ability in descriptive statistics. *Procedia-Social and Behavioral Sciences*, *129*, 133-139. https://doi.org/10.1016/j.sbspro.2014.03.658

[6] Erika, F., Prahani, B.K., Supardi, Z.A.I., & Tukiran. (2018). The development of metacognition-based learning media for the industrial electronics field in a vocational high school. World Trans. on Engng. and Technol. Educ., 16(2), 179-185.

[7] Iramaneerat, C. H. E. R. D. S. A. K., Smith Jr, E. V., & Smith, R. M. (2008). An introduction to Rasch measurement. *Best practices in quantitative methods*, 50-70. https://doi.org/10.4135/9781412995627.d6

[8] Jatmiko, B., Prahani, B.K., Munasir, Supardi, Z.A.I., Wicaksono, I., Erlina, N., Pandi-angan, P., Althaf, R., and Zainuddin. (2018). The comparison of OR-IPA teaching model and problem based learning model effectiveness to improve critical thinking skills of pre-service physics teachers. Journal of Baltic Science Education, 17(2), 1-22.

[9] Khan, W., & Ullah, H. (2010). Scientific Reasoning: A Solution to the Problem of Induction. *International Journal of Basic & Applied Sciences*, *10*(3), 58-62.

[10] Kuhn, D., & Franklin, S. (2006). *The second decade: What develops (and how)*. John Wiley & Sons, Inc..

[11] Limatahu I., Suyatno, Wasis, and Prahani, B.K., The effectiveness of CCDSR learning model to improve skills of creating lesson plan and worksheet science process skills (SPS) for pre-service physics teacher. J. Phys. Conf. Ser., 997, 32, 1-7 (2018). https://doi.org/10.1088/1742-6596/997/1/012032

[12] Prahani, B.K., Limatahu, I., Winata, S.W., Yuanita, L., & Nur, M. (2016). Effectiveness of physics learning material through guided inquiry model to improve student's problem solving skills based on multiple representation. International Journal of Education and Research. 4 (12), 231-244.

[13] Prahani, B.K., Nur, M., Yuanita, L., & Limatahu, I. (2016). Validitas model pembelajaran group science learning: Pembelajaran inovatif di Indonesia [Validity of learning model of group science learning: Innovative learning in Indonesia]. Vidhya Karya, 31(1), 72-80. https://doi.org/10.20527/jvk.v31i1.3976

[14] Prahani, B.K., Suprapto, N., Suliyanah, Lestari, N.A., Jauhariyah, M.N.R, Admoko, S., and Wahyuni, S., (2018). The effectiveness of collaborative problem based physics learning (CPBPL) model to improve student's self-confidence on physics learning. Journal Physics: Conference Series, 997(08), 1-6. https://doi.org/10.1088/1742-6596/997/1/012008

[15] Rangkuti, A. N. (2017). TANTANGAN DAN PELUANG PEMBELAJARAN MATE-MATIKA. *LOGARITMA: Jurnal Ilmu-ilmu Kependidikan dan Sains*, *2*(1), 1-13. https://doi.org/10.24952/logaritma.v5i01.1257

[16] Rizta, A., Zulkardi, Z., & Hartono, Y. (2013). Pengembangan Soal Penalaran Model TIMSS Matematika SMP. *Jurnal Penelitian dan Evaluasi Pendidikan*,*17*(2), 230-240. https://doi.org/10.21831/pep.v17i2.1697

[17] Salz, S. (2009). *Take the Test: Sample Questions from OECD's PISA Assessments*. OECD Publishing. 2, rue Andre Pascal, F-75775 Paris Cedex 16, France. https://doi.org/10.1177/0047287594033002101

[18] Siregar, H. (2003). Peranan Fisika Pada Disiplin Ilmu Teknik Kimia. *Sumatera Utara: USU digital library*.

[19] Stephens, A. L. & Clement, J.J. 2010. Documenting The Use of Expert Scientific Reasoning Processes by High School Physics Students. Physical Review Special Topics-Physics Education Research. 6(2),020122: 1-15. https://doi.org/10.1103/physrevstper.6.020122

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

[21] Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan rasch pada assessment pendidikan*. Trim Komunikata.

[22] Sunarti T., Wasis, Madlazim, Suyidno, and Prahani, B.K. (2018). The effectiveness of CPI model to improve positive attitude toward science (PATS) for pre-service physics teacher. Journal Physics: Conference Series, 997(13), 1-7. https://doi.org/10.1088/1742-6596/997/1/012013

[23] Suyidno, Nur, M., Yuanita, L., Prahani, B.K., and Jatmiko, B. (2018). Effectiveness of creative responsibility based teaching (CRBT) model on basic physics learning to increase

student's scientific creativity and responsibility. Journal of Baltic Science Education, 17(1), 136-151. https://doi.org/10.9790/7388-0701025661

[24] Thompson, E. D., Bowling, B. V., & Markle, R. E. (2018). Predicting student success in a major's introductory biology course via logistic regression analysis of scientific reasoning ability and mathematics scores. *Research in Science Education*, *48*(1), 151-163. https://doi.org/10.1007/s11165-016-9563-5

[25] Wati, M., Hartini, S., Hikmah, N., & Mahtari, S. (2018, March). Developing physics learning media using 3D cartoon. In Journal of Physics Conference Series (Vol. 997, No. 1). https://doi.org/10.1088/1742-6596/997/1/012044

[26] Zhou, S., Han, J., Koenig, K., Raplinger, A., Pi, Y., Li, D., ... & Bao, L. (2016). Assessment of scientific reasoning: The effects of task context, data, and design on student reasoning in control of variables. Thinking skills and creativity, 19, 175-187. https://doi.org/10.1016/j.tsc.2015.11.004

6 Authors

Mustika Wati is a senior lecturer at Physics Education Study Program. She holds a doctors degree in education research and evaluation. She is a research area of focus is physics education. (mustika_pfis@ulm.ac.id)

Saiyidah Mahtari is a lecture at Physics Education Study Program, FKIP Universitas Lambung Mangkurat. She holds a masters degree in Science Education. She is a research area of focus is physics education.

Sri Hartini is a senior lecturer at Physics Education Study Program. She holds a masters degree in Physics. She is a research area of focus is physics education. (srihartini_pfis@ulm.ac.id)

Heny Amalia is a college student at Physics Education Study Program.