ISSN 2964-4291
https://proceeding.unnes.ac.id/index.php/iset

Analysis of Junior High School Students' Scientific Literacy Using the Rasch Model

Mutiara Nurul Lita Azizah[1]*, Sri Ngabekti[2], Sigit Saptono[2], Endang Susilaningsih[2]

[1]Doctoral Program of Science Education, Universitas Negeri Semarang

[2] Universitas Negeri Semarang, Indonesia

*Corresponding Author: mutiaranla2410@students.unnes.ac.id

Abstract. This study aims to analyze the essay test instrument with 11 questions using the Rasch model to measure scientific literacy skills in the matter of Substances, Forms of Substances, and Their Changes. This research is a type of quantitative descriptive research. The research subjects were 84 class VII students at SMP PGRI 1 Ajibarang. The research data were analyzed using the winstep program to obtain a logit value. From the output of the Winsteps program, the results obtained were 9 questions according to the Rasch model with an average ± average Outfit MNSQ for person and item ± 0.98 and 0.98 respectively. Meanwhile, the Outfit ZSTD values for persons and items are ± 0 and -0.01, respectively. While the reliability of the instrument expressed in Cronbach's alpha is 0.95.

Keywords: junior high school; scientific literacy; the Rasch model.

INTRODUCTION

The quality of education is inseparable from the assessment system and the quality of learning that supports it. A good learning system will produce good quality learning (Mardapi, 2016). Education is the main key to the progress of the nation, and efforts to improve the quality of education must continue to be carried out so that Indonesia is able to compete in the era of globalization. Educators or teachers besides having to be able to evaluate the learning process, must also be able to develop evaluation tools or instruments used in the learning process according to the type of learning outcomes expected. In addition, evaluation tools or instruments must also be adapted to techniques for measuring and obtaining data which can then serve as indicators of achievement of learning objectives. Apart from that, the emergence of the impact of globalization requires the educational process to be able to produce students who have high quality so that they can survive and even become pioneers as citizens of the 21st century.

In addition to the learning approach, assessment also plays a role as an important component in learning. Assessment is a series of activities carried out by the teacher to obtain and process information about student learning processes and outcomes in a lesson (Nurlenasari et al., 2019). Assessment of student learning outcomes by educators should be carried out continuously. Assessment is not only intended to find out whether or not the goals that have been set have been achieved, but also to find out whether these goals are important to students and how students achieve them. Good assessment can be done by collecting accurate evidence regarding the achievement of student learning outcomes and making the class assessment process and the results beneficial to students, namely being able to increase their motivation and learning achievement (Stiggins & Chappuis, 2012: 3). The test is a way to assess the level of human ability indirectly, namely through a person's response to a number of stimuli or questions (Mardapi, 2008). A good quality test has the characteristics of good test items and devices.

Russel & Airasian argue that assessment is an important component in learning (Sumarnia, 2019). Assessment requires a tool to measure students' abilities. One of the tools used is a test, so that the tests carried out can find out the actual abilities of students, a test kit is needed. A good test set has three criteria, namely; the contents of the text are in accordance with the material to be tested (content validity), the text has a good construct (construct validity), and has constancy (reliability). The test is said to be reliable if it is used to measure several times, either for the same or different participants, the results are relatively the same.

A good assessment instrument must meet several criteria including good item validity, good item reliability, have varying levels of item difficulty, and have different power items that are able to distinguish students who are intelligent and capable of answering questions from students who are not. able to answer questions. The higher the value of the validity and reliability of an

instrument, the more precise the data obtained from a study (Hayati & Lailatussaadah, 2016), validity and reliability are also important factors in determining whether a test has good criteria or not (Tri Wahyuningsih, 2015). Analysis of test instruments in evaluation in the field of education can be done through two approaches. The first approach, which is the most widely and commonly applied approach in the field of education to date, is classical test theory (CTT). The aspects that greatly determine the quality of the items in the classical test theory approach are the level of difficulty and the discriminating power of the items. However, the characteristics of the items produced by the classical test theory are inconsistent (changing) depending on the ability of the test takers.

The second approach is a modern approach with Rasch modeling. Rasch modeling provides a different approach to the use of scores or raw test data in the context of educational assessment. The application of Rasch modeling measurements on the raw data of test results aims to produce a measurement scale with the same intervals which in turn can provide accurate information about the ability of test takers and the quality of the questions worked by students. Analysis of the items with Rasch modeling will produce information about the characteristics of the items and students that have been formed into the same metric (Sumintono & Widhiarso, 2015).

Winstep software is a computational tool for the Rasch model to analyze scores generated from test instruments with the aim of knowing Outfit MNSQ, Outfit ZSTD, Point Measure Correlation, Item reliability and Alpha Cornbach. The MNSQ outfit is useful for seeing the suitability of the data with the model used. The expected mean square value is 1 (one). If the mean-square value at infit is greater than one, the variation of the instrument is more than predicted by the Rasch model. If the infit value is less than 1, then the variation in the instrument is less when compared to the predictions made by the Rasch model.

In this study, the researcher wanted to know the quality of the scientific literacy test instrument developed by the researcher which was used to determine students' scientific literacy skills in the matter of Substances, Forms of Substances, the changes using the Rasch model approach. This quality is measured based on several indicators, namely the items that fit the Rasch model and the reliability of the items. Therefore a test instrument was designed and then it was determined which items were fit and which were not fit with the Rasch model. In addition, with the help of Winstep software, the Crobach alpha value will be determined to determine the reliability of the items.

METHODS

This study focused on the analysis of test instruments using the Rasch model. The subjects of this study were students of SMP PGRI 1 Ajibarang consisting of 85 students. There are 9 items on the scientific literacy multiple choice test instrument that are tested on students. The level of difficulty of the questions starts from easy, medium and difficult with relatively the same comparison. The questions are in the form of multiple choices with a correct score of 1 and an incorrect score of 0. So the data obtained is dichotomous data. The test results in the form of scores were analyzed using Winsteps software. From the output of the Winsteps software, several item parameters are obtained that fit the Rasch model. In addition, Cronbach's alpha value was obtained which was the result of the overall item reliability test. Meanwhile, Outfit MNSQ, Outfit ZSTD and item correlation values with questions as a whole indicate the limits of items that are declared fit with the model. That is, if the Outfit MNSQ value is between 0.5 and 1.5; Outfit ZSTD value is between -2.0 to 2.0; as well as item correlation values with a total score ranging from 0.4 to 0.85 (Sumintono & Widhiarso, 2015, p.12).

RESULTS AND DISCUSSION

Based on data analysis using Winsteps software, there are 11 items that fit the Rasch model. These results are presented in full in Table 1.

Tabel 1. Statistical Summary

	Description	Value
Logit	Person	0.75
	Item	0
Reliabilitas	Person Realibility	0.74
	Item Realibility	0.77
	Alpha Cronbach	0.85
Outfit MNSQ	Person	0.98
	Item	0.98
Outfit ZSTD	Person	0
	Item	-0.1

Table 1 shows the logit value of the person or measure of 0.75 and the value of the item measure of 0, which means that the value of the

person measure is greater than the item measure. It can be stated that students' abilities tend to be higher than the level of difficulty. In other words, there is a possibility that all items can be answered correctly by students. So that students who have the highest ability can answer the most difficult questions correctly. Meanwhile, item reliability is worth 0.77, Person Reliability is worth 0.74 and Cronbach's Alpha value is 0.85. From this value it can be stated that the level of consistency of students' answers is quite high, and the quality of the items in the test instrument used has a fairly good reliability, namely 0.77. besides that, the value of Cronbach's Alpha which shows the interaction between person and item as a whole is quite good, namely 0.85.

Another quantity shown in table 1 is the Outfit Mean Squared (Outfit MNSQ) value of 0.98 in both the person and item columns. The value of 0.98 is included in the fit criterion, which is located between the intervals of 0.5<MNSQ<1.5, meaning that the test instrument used is in accordance with the model to measure student competence in the material for official life insurance calculations. Furthermore, Outfit Z Standardized (Outfit ZSTD) values are 0 for persons and -0.01 for items. The values 0 and -0.01 are between the interval -2.0 <ZSTD < 2.0 which can be interpreted as having a rational possible value. This means that as a whole the items or items are in accordance with the Rasch model and can be used as an achievement test instrument in the material for calculating life insurance premiums.

The distribution of item items considered to be misfit or not fit to the model can be seen in Table 2. Item limits are declared fit to the model if one or both of the following conditions are met. The first requirement is that the Outfit MNSQ value lies between 0.5 to 1.5; Outfit ZSTD value lies between -2.0 to 2.0; and the correlation value of items with a total score (point measure correlation) lies between 0.4 to 0.85 (Sumintono & Widhiarso, 2014, pp. 71-72)

Table 2. Data on the distribution of questions that are misfit or not fit with the Rasch model

Person STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Person
3	10	11	2.84	1.10	1.09	.37	.68	.21	.26	.28	90.9	90.9	03
16	9	11	1.93	.85	.84	-.22	.55	-.19	.52	.38	90.9	81.9	16
18	9	11	1.93	.85	.84	-.22	.55	-.19	.52	.38	90.9	81.9	18
23	9	11	1.93	.85	1.17	.50	.75	.08	.33	.38	72.7	81.9	23
5	8	11	1.29	.76	1.01	.16	.75	-.15	.47	.44	72.7	76.8	05
15	8	11	1.29	.76	1.17	.57	1.29	.61	.30	.44	72.7	76.8	15
20	8	11	1.29	.76	1.01	.16	.75	-.15	.47	.44	72.7	76.8	20
26	8	11	1.29	.76	.98	.06	.73	-.19	.49	.44	72.7	76.8	26
34	8	11	1.29	.76	.70	-.83	.52	-.60	.67	.44	90.9	76.8	34
35	8	11	1.29	.76	1.01	.16	.75	-.15	.47	.44	72.7	76.8	35
36	8	11	1.29	.76	.70	-.83	.52	-.60	.67	.44	90.9	76.8	36
44	8	11	1.29	.76	1.36	1.03	1.30	.63	.20	.44	72.7	76.8	44
46	8	11	1.29	.76	.90	-.16	.67	-.29	.54	.44	72.7	76.8	46
42	7	11	.74	.72	.53	-1.66	.43	-1.25	.81	.48	90.9	73.1	42
53	7	11	.74	.72	1.01	.15	.89	-.04	.49	.48	72.7	73.1	53
57	7	11	.74	.72	1.23	.77	1.13	.41	.34	.48	72.7	73.1	57
1	6	11	.24	.70	.60	-1.37	.51	-1.28	.79	.50	90.9	72.5	01
8	6	11	.24	.70	.90	-.21	.80	-.37	.59	.50	72.7	72.5	08
10	6	11	.24	.70	1.53	1.54	1.57	1.29	.13	.50	54.5	72.5	10
19	6	11	.24	.70	.92	-.16	.83	-.30	.57	.50	72.7	72.5	19
22	6	11	.24	.70	1.06	.29	1.05	.27	.46	.50	72.7	72.5	22
27	6	11	.24	.70	1.00	.10	.93	-.04	.51	.50	72.7	72.5	27
28	6	11	.24	.70	.90	-.21	.80	-.37	.59	.50	72.7	72.5	28
29	6	11	.24	.70	.90	-.21	.80	-.37	.59	.50	72.7	72.5	29
31	6	11	.24	.70	.60	-1.37	.51	-1.28	.79	.50	90.9	72.5	31
33	6	11	.24	.70	.77	-.68	.74	-.55	.66	.50	90.9	72.5	33
41	6	11	.24	.70	2.13	2.82	2.75	3.00	-.36	.50	36.4	72.5	41
43	6	11	.24	.70	.77	-.68	.74	-.55	.66	.50	90.9	72.5	43
49	6	11	.24	.70	.60	-1.37	.51	-1.28	.79	.50	90.9	72.5	49
50	6	11	.24	.70	.77	-.68	.74	-.55	.66	.50	90.9	72.5	50
55	6	11	.24	.70	.77	-.68	.74	-.55	.66	.50	90.9	72.5	55

58	6	11	.24	.70	.94	-.10	.87	-.20	.56	.50	72.7	72.5	58
63	6	11	.24	.70	.98	.04	.84	-.26	.54	.50	72.7	72.5	63
68	6	11	.24	.70	1.11	.44	.98	.10	.45	.50	54.5	72.5	68
73	6	11	.24	.70	.83	-.47	.78	-.42	.62	.50	72.7	72.5	73
74	6	11	.24	.70	1.11	.44	.98	.10	.45	.50	54.5	72.5	74
75	6	11	.24	.70	1.51	1.49	1.95	1.91	.09	.50	54.5	72.5	75
4	5	11	-.24	.70	.85	-.40	.79	-.40	.61	.50	81.8	73.6	04
6	5	11	-.24	.70	.43	-2.21	.38	-1.82	.89	.50	100.0	73.6	06
9	5	11	-.24	.70	1.34	1.05	1.44	1.06	.25	.50	63.6	73.6	09
30	5	11	-.24	.70	.43	-2.21	.38	-1.82	.89	.50	100.0	73.6	30
32	5	11	-.24	.70	.43	-2.21	.38	-1.82	.89	.50	100.0	73.6	32
38	5	11	-.24	.70	.43	-2.21	.38	-1.82	.89	.50	100.0	73.6	38
39	5	11	-.24	.70	.43	-2.21	.38	-1.82	.89	.50	100.0	73.6	39
40	5	11	-.24	.70	1.46	1.36	1.55	1.26	.17	.50	63.6	73.6	40
51	5	11	-.24	.70	.43	-2.21	.38	-1.82	.89	.50	100.0	73.6	51
52	5	11	-.24	.70	.69	-1.00	.59	-1.01	.73	.50	81.8	73.6	52
59	5	11	-.24	.70	1.16	.57	1.03	.21	.42	.50	63.6	73.6	59
64	5	11	-.24	.70	1.37	1.15	1.18	.55	.29	.50	45.5	73.6	64
66	5	11	-.24	.70	1.11	.43	1.00	.15	.45	.50	63.6	73.6	66
67	5	11	-.24	.70	.78	-.62	.69	-.70	.66	.50	81.8	73.6	67
72	5	11	-.24	.70	.79	-.60	.74	-.54	.65	.50	81.8	73.6	72
76	5	11	-.24	.70	1.53	1.53	1.41	1.00	.17	.50	45.5	73.6	76
83	5	11	-.24	.70	1.00	.11	.86	-.22	.53	.50	63.6	73.6	83
84	5	11	-.24	.70	.74	-.79	.66	-.80	.69	.50	81.8	73.6	84
21	4	11	-.74	.72	.95	-.04	.97	.12	.51	.48	81.8	73.4	21
24	4	11	-.74	.72	.57	-1.46	.46	-1.15	.79	.48	81.8	73.4	24
45	4	11	-.74	.72	.69	-.97	.55	-.87	.71	.48	81.8	73.4	45
60	4	11	-.74	.72	1.70	1.90	2.44	2.18	-.08	.48	45.5	73.4	60
61	4	11	-.74	.72	1.70	1.90	2.44	2.18	-.08	.48	45.5	73.4	61
70	4	11	-.74	.72	.57	-1.46	.46	-1.15	.79	.48	81.8	73.4	70
81	4	11	-.74	.72	.83	-.44	.69	-.50	.62	.48	81.8	73.4	81
82	4	11	-.74	.72	.62	-1.24	.50	-1.03	.76	.48	81.8	73.4	82
2	3	11	-1.29	.76	1.31	.93	2.88	2.04	.08	.44	72.7	76.1	02
13	3	11	-1.29	.76	1.32	.94	1.30	.63	.22	.44	72.7	76.1	13
17	3	11	-1.29	.76	1.58	1.51	1.48	.83	.06	.44	54.5	76.1	17
25	3	11	-1.29	.76	1.32	.94	1.15	.44	.25	.44	54.5	76.1	25
47	3	11	-1.29	.76	.58	-1.29	.42	-.84	.75	.44	90.9	76.1	47
54	3	11	-1.29	.76	1.50	1.34	2.25	1.55	.02	.44	54.5	76.1	54
56	3	11	-1.29	.76	.96	.00	1.03	.28	.45	.44	90.9	76.1	56
69	3	11	-1.29	.76	1.39	1.09	1.20	.50	.21	.44	54.5	76.1	69
71	3	11	-1.29	.76	1.38	1.09	1.67	1.03	.13	.44	72.7	76.1	71
77	3	11	-1.29	.76	1.03	.20	.88	.05	.44	.44	72.7	76.1	77
78	3	11	-1.29	.76	1.03	.20	.88	.05	.44	.44	72.7	76.1	78
79	3	11	-1.29	.76	1.03	.20	.88	.05	.44	.44	72.7	76.1	79
80	3	11	-1.29	.76	.89	-.19	.79	-.09	.52	.44	90.9	76.1	80
7	2	11	-1.93	.85	.72	-.53	.44	-.36	.60	.38	90.9	82.1	07
37	2	11	-1.93	.85	1.72	1.46	4.04	2.08	-.35	.38	72.7	82.1	37
48	2	11	-1.93	.85	1.33	.81	1.26	.58	.16	.38	72.7	82.1	48
62	2	11	-1.93	.85	1.24	.64	1.20	.52	.21	.38	72.7	82.1	62
85	2	11	-1.93	.85	1.80	1.58	6.37	2.86	-.57	.38	72.7	82.1	85
14	1	11	-2.84	1.10	.69	-.20	.28	-.26	.52	.28	90.9	90.9	14
65	1	11	-2.84	1.10	1.12	.41	.75	.27	.23	.28	90.9	90.9	65
11	0	11	-4.19	1.88	MINIMUM MEASURE				.00	.00	100.0	100.0	11
12	0	11	-4.19	1.88	MINIMUM MEASURE				.00	.00	100.0	100.0	12
MEAN	5.0	11.0	-.27	.77	1.00	.0	1.04	.0			76.0	75.5	
P.SD	2.1	.0	1.22	.19	.36	1.1	.86	1.0			14.8	4.1	

Based on the results of the analysis of the achievement instrument using the help of the Winsteps version 4.4.5 program in the table, it was obtained 25 items that were misfit, and 25 items that were fit, so that 25 items were obtained as a final instrument.

The value of the student's ability level in working on the questions is shown from the winsteps output, namely the Wright map. Data obtained from students with codes L13, L14, L17, P02, P27 and P40 have the highest ability or ability, while student P39 has the lowest ability.

CONCLUSION

The test instrument used for actuarial

mathematics courses on material for calculating fit life insurance premiums with the Rasch model. This is indicated by the item score (item reliability) of 0.77, person reliability of 0.74, and Cronbach's alpha value of 0.85 while the Outfit Mean Square Statistics (Outfitt MNSQ) value is 0.98 in the person and item column. The Outfit Z Standard (Outfit ZSTD) value is 0 for the person table and -0.01 for the item table. While the number of items that fit were 25 while those that did not fit were 25 items

REFERENCES

Hayati, S., & Lailatussaadah, L. (2016). Validitas dan Reliabilitas Instrumen Pengetahuan Pembelajaran Aktif, Kreatif dan Menyenangkan (Pakem) Menggunakan Model RASCH. Jurnal Ilmiah Didaktika, 16(2), 169.

Mardapi, D. (2008a). Teknik penyusunan instrumen tes dan non tes. Yogyakarta: Mitra Cendekia Offset.

Mardapi, D. (2016b). *Pengukuran, penilaian dan evaluasi pendidikan edisi 2*. Yogyakarta: Nuha Litera

Nurlenasari, N., Lidinillah, D. A. M., Nugraha, A., & Hamdu, G. (2019). Assessing 21st century skills of fourth-grade student in STEM learning. Journal of Physics: Conference Series, 1318(1)

Sri Sumarnia, "Designing Ict Competences-Integrated Assessment Instruments Of Practical Key Teaching Competences For English Language Education Study Program," Ijlec International Journal Of Language Education And Culture Review 5, no. 1 (2019): 47–55

Stiggins, R.J. & Chappuis, J. (2012). *An introduction to student involved assessment for learning (6th ed.*). Boston: Pearson.

Sumintono, B., & Widhiarso, W. (2015). Aplikasi Pemodelan Rasch pada Assessment Pendidikan. Cimahi: Trim Komunikata.

Tri Wahyuningsih, E. (2015). Analisis Butir Soal Tes Objektif Buatan Guru Ulangan Semester Ganjil Mata Pelajaran Ekonomi Kelas X Di Sma Negeri 1 Mlati Tahun Ajaran 2013/2014 (Universitas Negeri Yogyakarta). Retrieved from http://eprints.uny.ac.id/id/eprint/26627.