International Journal of Education, Information Technology and Others (IJEIT), August 2023, 6 (3), 521-533
DOI: https://doi.org/10.5281/zenodo.10183222
p-ISSN: 2654-2528 e-ISSN: 2623-2324
Accredited by Directorate General of Strengthening for Research and Development
Available online at https://jurnal.peneliti.net/index.php/IJEIT

Analysis of Characteristics of Final Semester Examination Questions for Class IX Mathematics Using Rasch Model Theory in the Tidore Islands Region

Aji Joko Budi Pramono[1], Yani Jawa[2]

[1,2] Lecturer at Ternate State Islamic Institute

Received: 8 July 2023
Revised: 12 July 2023
Accepted: 29 July 2023

Abstract

This study aims to analyze the characteristics of grade IX mathematics exam questions in the Tidore Islands region, North Maluku, using a quantitative descriptive methodology approach. Involving ninth grade junior high school students in North Maluku, this study used the Rasch Model approach based on the R program to analyze the quality of the items, including the unidimensional assumption test, item difficulty test, student ability test and item information value test. The instrument used was a test form from the mathematics subject. The results showed that the application of item response theory in analyzing Mathematics Final Semester Examination data was successful, with 11 factors and 1 dominant factor supporting the assumption of unidimensionality. The focus on the difficulty level of the questions showed that most of the questions were in the good range, effective in differentiating students' abilities. Analysis of the item characteristic curves provided insight into the quality of item design, while the accuracy of these curves, as per the Rasch model, is key in improving measurement validity. Thus, the results of this study are significant in improving understanding of the characteristics of grade IX mathematics items in the tidore islands region.

Keywords: Rasch Model, Item Characteristics, Validity

(*) Corresponding Author: ajijoko@gmail.com
How to Cite: Pramono, A. J., & Jawa, Y. (2023). Analysis of Characteristics of Final Semester Examination Questions for Class IX Mathematics Using Rasch Model Theory in the Tidore Islands Region. *International Journal of Education, Information Technology, and Others*, *6*(3), 521-533. https://doi.org/10.5281/zenodo.10183222

INTRODUCTION

Education has an important role in individual growth and societal progress. Teachers have a central role in facilitating effective learning. In an effort to improve learning, assessment and evaluation are key. However, it is important to understand the quality of the questions in creating quality evaluation tools, such as tests or exams. The purpose of the assessment carried out by the teacher is of course to see the progress of the learning process and obtain information about students' progress (Tiara & Sari, 2019). Evaluation of the learning process needs to be carried out to find out students' actual abilities (Zainal, 2020). A good quality assessment model will certainly require a form of examination instrument that has good quality (Sanjaya Putra & Renda, 2022). An instrument is said to have good quality if the instrument is able to measure test participants validly (Sanjaya Putra & Renda, 2022). An instrument that cannot contribute any information to the test participants being measured means the instrument has no validity value (Zohrabi, 2013). The

large number of instruments that do not meet the validity criteria will cause students' abilities to not be measured properly (Mustafa & Masgumelar, 2022), thus requiring quality questions. Good questions must be valid and relevant, accurately measuring students' abilities in certain subjects. Therefore, analyzing the quality of the questions is very important.

There are two models for analyzing question items, namely using the classical model and modern theory. The modern theoretical approach is to use the Rasch Model, which is based on Item Response Theory (TRB) (Kurnia, 2019). Classical theory is starting to be abandoned by academics, because it has several shortcomings, especially in analyzing individual test items (Pratama, 2020). The weaknesses of classical theory include, among other things, the characteristics of test equipment will influence measurement results, ability parameters and measurements based on group tests, not based on tests. individual (Djemari, 2017, p. 17). The weakness of this classical theory can be corrected by using TRB, this theory directly compares the test taker's student item responses with the student's individual abilities (Djemari, 2017, p. 24) TRB always assumes that the chance that the test taker will answer correctly for each question really depends on his own ability, in other words students who have high ability will easily answer easy questions than students who have low ability (Retnawati, 2014)

The problem that arises in the quality of test items is how to improve the quality of test items effectively using the Rasch Model, so that education as a whole can improve, and assessment instruments can be in line with learning objectives, provide reliable results, and support better student learning outcomes. This problem is related to the validity of the test items, because this will ensure that the test items can measure students' learning outcomes and abilities accurately. In addition, we need to understand the level of difficulty of the items so that we can adjust teaching strategies appropriately according to students' needs. In addition to the validity and difficulty of the test items, it is also important to assess the information value of the test items. This will help us determine the extent to which the items can differentiate students' abilities well.

The importance of analyzing the quality of questions using the Rasch Model can be seen from several aspects. First, the Rasch Model allows measuring the level of difficulty of the items. In other words, teachers can find out to what extent a question item is considered difficult or easy by students. This information is very valuable because it allows teachers to adjust learning to suit the expected level of item difficulty. Second, the Rasch Model provides the information value of an item, this describes how well an item is at identifying differences between students with high and low ability. Question items with high information value will make a greater contribution in differentiating between students with different abilities. Third, the Rasch Model ensures a match between students' abilities and the expected model. In this context, the desired model is the Rasch model.

A very important novelty in this research is the emphasis on increasing the validity of the item instrument. Validity is an essential basis for creating effective evaluation tools. The use of the Rasch Model in analyzing test items makes a significant contribution in ensuring that the assessment instrument is able to measure student learning outcomes and abilities accurately. By using the Rasch Model, this research presents the ability to in-depth and accurately evaluate each

question item individually, measuring the level of difficulty, validity and information value of each question item, which ultimately provides greater confidence that the evaluation results reflect actual achievement. of students in a particular subject and is not influenced by irrelevant factors. In comparison with the more commonly used classical methods previously, the Rasch Model creates a significant step forward in improving the accuracy and objectivity of item assessment, bringing real benefits to student learning and the development of the overall quality of educational assessment. By prioritizing the validity of the test item instrument, this research brings critical innovation to learning and educational evaluation. This provides assurance that the evaluation tools used by educators and educational institutions can provide accurate and reliable results, which will ultimately support better student learning and improve the overall quality of education. This novelty creates a bridge between theory and practice, allowing educators to take concrete steps in improving the validity of their item instruments. Thus, this research provides significant added value in efforts to improve the quality of assessment and learning in the educational context

The importance of analyzing the quality of test items using the Rasch Model also lies in efforts to improve the overall quality of education. By using this tool, teachers can more effectively identify question items that are poor and need to be corrected or replaced. This will have a positive impact on student learning outcomes and ultimately, on educational progress at large. In order to improve the quality of assessment and evaluation instruments in education, it is important for teachers and educational decision makers to understand and implement analysis of the quality of test items using the Rasch Model. In this way, we can ensure that the evaluation tools used are appropriate to the learning objectives and provide accurate results to guide better learning and improve student abilities.

METHOD

This research is a scientific effort that uses a quantitative approach to understand the characteristics of class IX mathematics exam questions in the Tidore Islands region, North Maluku. The methodology used in this research is quantitative descriptive. The researcher focused the analysis on the parameters of the questions using the R program and the Winstep program, with the method*Rasc Model*. This approach makes it possible to identify questions that are in accordance with the Rasc Model concept, while questions that are not appropriate will be eliminated from the research instrument.

This research was carried out in North Maluku and involved class IX SMP/MTs students in the area. The research implementation period took place from May to November 2023. The data used came from UAS test responses given to respondents who were class IX students in the Tidore area, North Maluku. The data collection process was carried out using instruments that had been methodologically tested. Before the instrument was distributed to respondents, the suitability of the instrument was tested using the Rash Model test method. The research population consisted of class IX SMP/MTs students in the North Maluku province, while the research sample consisted of 50 respondents. Data analysis was carried out by testing the quality of the question items based on item response theory

tests*Rasc Model*. The initial step is to test the unidimensional assumption to ensure the characteristics of the test items. Then, the local independence assumption test is used to ensure that student responses and question items do not influence each other. Next, the test of the assumption of invariance of item parameters is used to assess the consistency of item characteristics in different groups of examinees. The information function of the items is used to measure the contribution of the items to the latent trait being measured. First, a unidimensional assumption test is carried out by looking at the KMO-MSA value, this value is considered sufficient if it is more than 0.5, apart from that, looking at the eigenvalue, factors with an eigenvalue more than 1 are the factors that are used.

DISCUSSION RESULTS

Results

Assumption Test Analysis

Data analysis using item response theory, the first thing that is done is to test the dimensions of the empirical data obtained. The testing process is done with an exploratory factor analysis method*main make up*. The initial assumptions that must be met in exploratory factor analysis are the KMO and Bartlett tests as follows.

Table 1. KMO and Bartlett

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.733
Bartlett's Test of Sphericity	Approx. Chi-Square	1233.700
	df	435
	Sig.	.000

Referring to Table 1 shows that in this research, it can be said to have sufficient data as indicated by KMO > 0.5. Apart from that, the significance of Bartlett's test shows that the Null Hypothesis of the correlation matrix is a dilected identity matrix so that the data forms a correlation matrix with a close relationship between variables.

In table 2,*Total Variance Explained* above shows that there are 11 factors formed from the 30 variables entered. Each factor's eigenvalue is > 1. Thus, empirical data shows that the mathematics UAN question items are divided into 11 factors. The results of factor analysis in Table 1 show that there are 11*eigenvalue* which has a value of more than 1, where the first factor is the most dominant factor. This factor has*eigenvalue* amounting to 4.433, the second to 11th dominant factor has*eigenvalue* more than 1.00

Table 2. Nilai Eigen

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	4.433	14.778	14.778
2	2.036	6.788	21.566
3	1.710	5.700	27.266
4	1.480	4.935	32.201
5	1.388	4.627	36.828
6	1.278	4.260	41.089
7	1.260	4.200	45.289
8	1.199	3.998	49.287
9	1.128	3.759	53.047
10	1.031	3.435	56.482
11	1.010	3.366	59.848
12	.938	3.127	62.975
13	.905	3.017	65.992
14	.865	2.883	68.875
15	.818	2.727	71.602
16	.784	2.615	74.217
17	.758	2.527	76.744
18	.738	2.460	79.205
19	.682	2.273	81.478
20	.643	2.145	83.623
21	.639	2.128	85.751
22	.605	2.017	87.768
23	.579	1.931	89.699
24	.546	1.821	91.520
25	.497	1.658	93.178
26	.470	1.567	94.745
27	.450	1.498	96.244
28	.415	1.382	97.625
29	.361	1.202	98.827
30	.352	1.173	100.000

Furthermore*eigenvalue* 12th to 30th have*eigenvalue* less than 1.00 this magnitude is considered the same. In factor analysis,*eigenvalue* the first one should have the greatest (dominant) value compared*eigenvalue* the second, third, and so on (Nasution, 2019). Because the amount of variance explained is directly proportional to the size*eigenvalue*, it can be concluded that the first factor in the factor analysis provides the greatest contribution compared to the other factors, so that the unidimensionality assumption is met. According to Hambleton & Rovinelli (1986) as quoted by Javili et all (2015) which states that usually the number of significant factors is determined by the appearance of an "elbow in the plot", the number of eigenvalues to the left of this elbow is interpreted as the number of dimensions formed. The scree plot in Figure 2 shows that an elbow has formed, with three points to the left of the elbow. By referring to the opinions of the experts above, it can be concluded that there are 11 factors formed in this instrument with 1 dominant factor

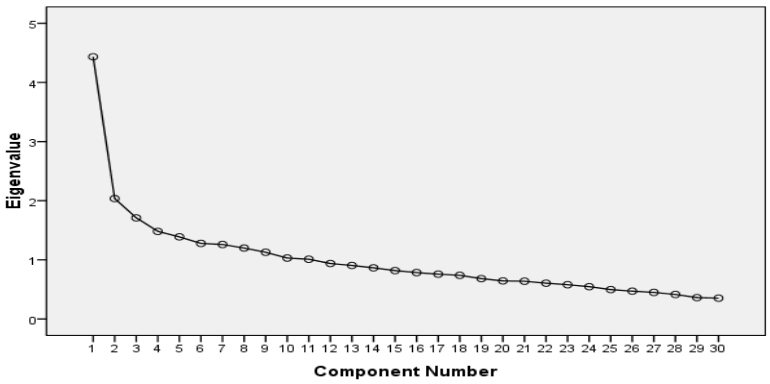


Figure 2. *Screeplot* explains the relationship between the number of factors formed and the eigenvalue

Characteristics of Question Items

Table 3 below shows the results of the analysis of the level of difficulty of the questions:

Table 3. Difficulty level of questions

No. Soal	Tingkat Kesulitan	Keteranga	No. Soal	Tingkat Kesulitan	Keteranga
x1	-5.05640615	Sangat Mudah	x13	-0.20029484	Sedang
x5	-1.28464639	Mudah	x27	-0.20004803	Sedang
x4	-1.15992329	Mudah	x7	-0.13646825	Sedang
x2	-1.13546437	Mudah	x18	-0.13639548	Sedang
x3	-1.08697718	Mudah	x22	-0.11523631	Sedang
x6	-0.83017275	Sedang	x25	-0.07316143	Sedang
x30	-0.80764169	Sedang	x8	-0.07272266	Sedang
x12	-0.58587537	Sedang	x24	0.01188927	Sedang
x9	-0.58577764	Sedang	x19	0.05420211	Sedang
x28	-0.52063713	Sedang	x15	0.0544224	Sedang
x23	-0.5205594	Sedang	x20	0.3110193	Sedang
x29	-0.47713252	Sedang	x21	2.12843128	Sangat sulit
x11	-0.45577441	Sedang	x10	2.45783694	Sangat sulit
x16	-0.36963856	Sedang	x17	5.18659579	Sangat sulit
x14	-0.28511176	Sedang	x26	5.52756187	Sangat sulit

Of the thirty items, there are twenty items that have a level of difficulty between -2 to +2, which means that the level of difficulty of the items is in the good range, in this range the level of difficulty is divided into a medium level of difficulty from -1 to 1, an easy level of difficulty from -2 to -1 and the difficulty level is difficult from 1 to 2. There is 1 question item which is classified as very easy, namely with a difficulty level range of more than -2 and there are 4 questions which are classified as very difficult, namely in the difficulty level range of more than -2. Of the 2, questions in the very easy category have a relatively small chance of students being able to answer them correctly, while questions in the very difficult category have relatively small chances of students answering correctly. Meanwhile, the item characteristic curves for the 30 questions are shown in Figure 3 as follows.

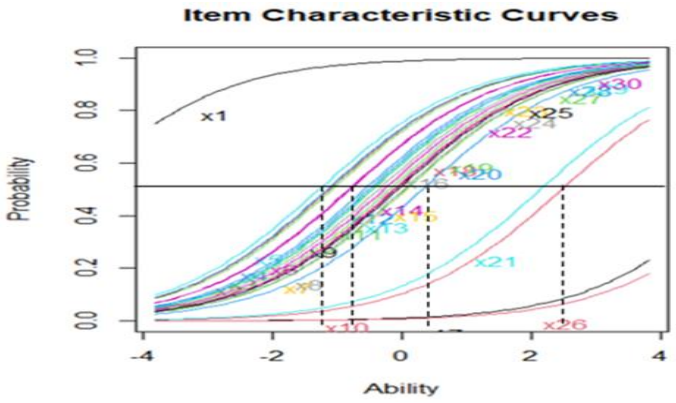


Figure 3. Characteristic curve for 30 questions

Based on the ICC graph for the 30 items, it appears that the level of difficulty in the very difficult category, namely items x21, x10, x17 and x26, has a graph that does not match the rascmodel graph, as well as the level of difficulty in the very easy category, namely item x1, has a graph. which does not match the pattern*Hive Model*. The easy, medium and difficult difficulty level categories have graphs according to the pattern*Rach Model*. KThe characteristic curve shows the relationship between ability and the probability of answering correctly. In theoretical analysis*Quick model* then the variable considered is the level of difficulty which is a point on the ability scale so that the chance of answering correctly is 0.5 (Retnawati, 2014) as in table 4.

Table 4. Difficulty Level of Question Items

No. Soal	Tingkat Kesulitan	Keterangan	No. Soal	Tingkat Kesulitan	Keterangan
x1	-5.05640615	Sangat Mudah	x13	-0.20029484	Sedang
x5	-1.28464639	Mudah	x27	-0.20004803	Sedang
x4	-1.15992329	Mudah	x7	-0.13646825	Sedang
x2	-1.13546437	Mudah	x18	-0.13639548	Sedang
x3	-1.08697718	Mudah	x22	-0.11523631	Sedang
x6	-0.83017275	Sedang	x25	-0.07316143	Sedang
x30	-0.80764169	Sedang	x8	-0.07272266	Sedang
x12	-0.58587537	Sedang	x24	0.01188927	Sedang
x9	-0.58577764	Sedang	x19	0.05420211	Sedang
x28	-0.52063713	Sedang	x15	0.0544224	Sedang
x23	-0.5205594	Sedang	x20	0.3110193	Sedang
x29	-0.47713252	Sedang	x21	2.12843128	Sangat sulit
x11	-0.45577441	Sedang	x10	2.45783694	Sangat sulit
x16	-0.36963856	Sedang	x17	5.18659579	Sangat sulit
x14	-0.28511176	Sedang	x26	5.52756187	Sangat sulit

Next, identification of the information function formed based on the data will be carried out. The information function of the item states the strength or contribution of the test item in revealing the latent trait being measured (Retnawati, 2014, p. 15). The following is the item information function that is formed based on data:

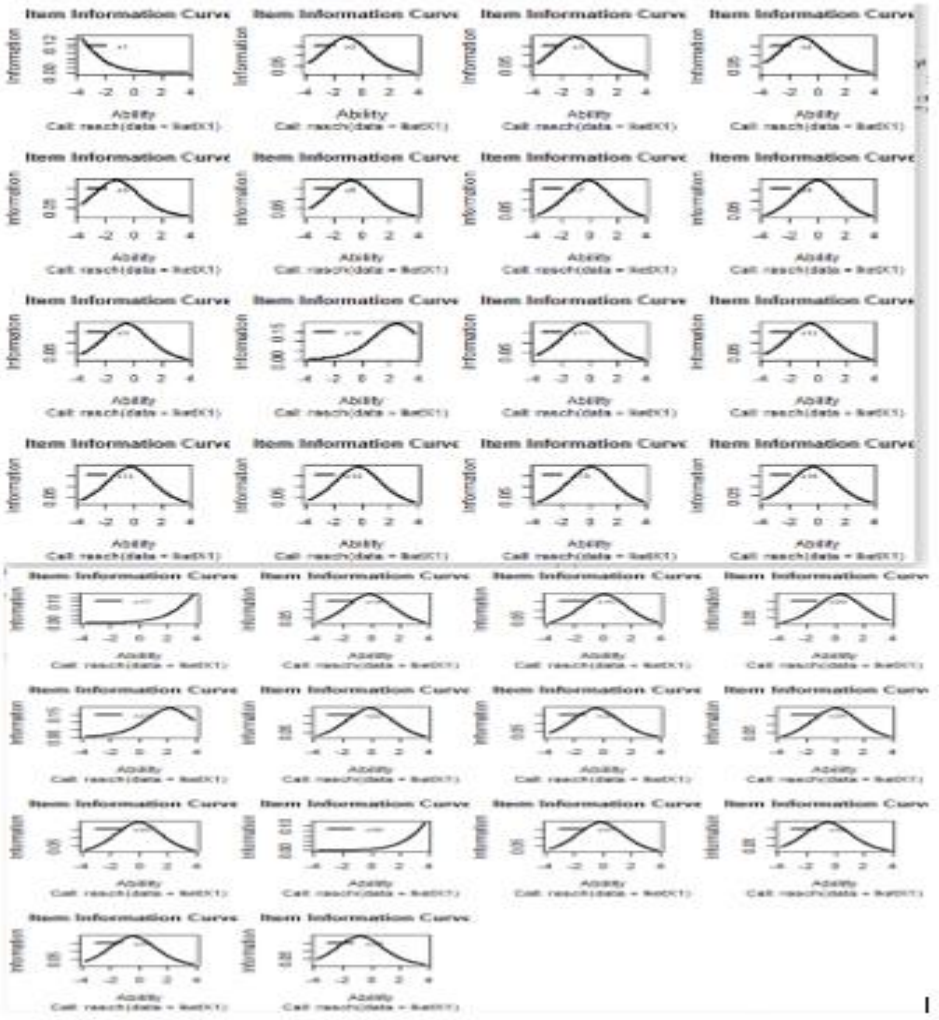


Figure 3. Graph of Information Value of Question Items

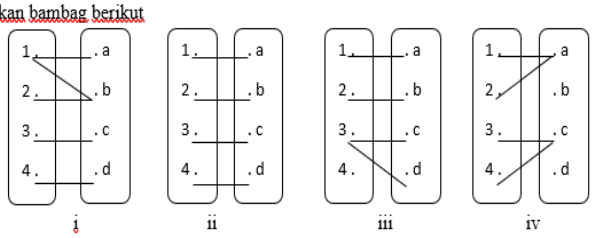
In a study of the analysis of item response theory and item characteristic curves, it was found that the difficulty level of the items tended to be at a medium level, indicating good quality of item design. Items at this level can effectively differentiate participants with different abilities. The graph in Figure 3 shows that items 1, 10, 17, and 26 have lower information value and indicate student abilities outside the range -2 to 2, so they may not be suitable for further analysis. The accuracy of the item characteristic curve according to the Rasch Model increases the validity of the measurement, essential for correct decisions in evaluating student abilities.

Descriptive analysis of question items:

1) Question Item Number 1

Based on theory*Quick model*, it is known that the difficulty level of question number 1 is -0.56, which indicates that the question is included in the very easy category

Pwrhatikan bambag berikut



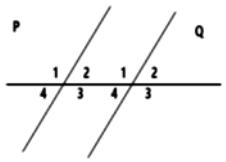
Relasi yang merupakan fungsi di tunjukkan pada gambar...

a. (i) dan (ii)
b. (i) dan (iii)
c. (ii) dan (iii)
d. (ii) dan (iv)

The Rasch model measures student ability in logit units, with item number 1 having a low level of difficulty (-0.56 logit), can be answered by students with very low ability (-0.4 logit), but many students answered correctly (239 of 250 students). The Rasch model curve graph does not match item number 1, and the information value of this item is 0.12 with student ability -0.4 logit, indicating that this item does not provide sufficient information. Conclusion, question number 1 has a low level of difficulty, provides little information, and needs improvement. Instructions are unclear, answer labels use Roman numerals, and questions only rely on diagrammatic interpretation without requiring a deeper understanding of mathematics.

2) Question Item Number 21

Based on the Rasch Model theory, item number 21 has a difficulty level of 2.2, which makes it included in the very difficult category. This can be seen from the student responses, where only 41 students out of a total of 250 students answered correctly. This means that more students answered incorrectly, indicating that item number 21 requires high student ability to answer correctly.

2 **1**	Perhatikan Gambar!  Diketahui besar sudut $P_1 = (2n+14)^0$ dan besar sudut $Q_3=(3n-30)^0$. Maka besar sudut Q_1 adalah a. 44^0 b. 78^0 c. 136^0 d. 102^0

Apart from that, the information function value of this item is low, only 0.14, so it has not made a significant contribution in revealing the latent trait being measured. This question item cannot differentiate between students with low and high ability in mastering latent traits.

3) Question Item Number 10

This question item has a difficulty level of 2.45, which indicates that the difficulty level of this question item is very difficult. This can be observed from the students' responses, where only 33 students out of a total of 250 students answered correctly. This means that most students answered incorrectly, which indicates that item number 10 requires high student ability to answer correctly.

10	A plane crashes and falls to the bottom of the sea. After being evacuated, 4 dive teams were assigned to search for the plane's black box. Based on evidence and information, it is known that the black box is at point A (5,-3). And the divers are spread out at several points. The first diver is at point P(3.5), the second diver is at point Q(-5.3), the third diver is at point R(-3.5), and the fourth diver is at point S(3,-5). If the 4 divers dive with a rotation direction of 90^0 clockwise, then the diver who finds the black box first is at the point... a. S b. R c. Q d. P

Question number 10 does not match the characteristics of the Rasch model, both in terms of the model itself and its graphical representation. The probability of answering question number 10 correctly does not match the assumptions*Quickly* Model. In addition, the information function of the items, which indicates their contribution in measuring hidden traits, is only 0.12, indicating a low information value. In conclusion, item 10 seems too difficult, requiring high student ability. However, these items do not match the characteristics of the Rasch model and provide minimal contribution in measuring hidden traits. Further evaluation is needed to improve the quality of the questions.

4) Question number 17

Item number 17 has a difficulty level of 5.2 on the logit scale, which indicates that this item is very difficult. This can be seen from the student responses, where only 6 out of 250 students answered correctly. This means that more students answer incorrectly, so that to answer question number 17 correctly, very high student ability is required.

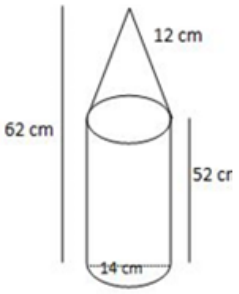
17	Dian filled the bathtub with a flow rate of 50 liters/minute and it was completely filled within 15 minutes. This water is used for various purposes. The next day the bathtub was empty and Dian filled it again with water, but it only filled up within 25 minutes. The water discharge at that time was... a. 20 liters/minute b. 30 liters/minute c. 40 liters/minute d. 50 liters/minute

If we look at the characteristics of the items according to the Rasch model, it can be seen that the characteristics of item number 17 do not match*Quick model*, both in terms of the model itself and the graphics. The probability of answering correctly is not consistent with the student's ability level. Furthermore, if the information value of a question item is identified based on data, the information function of the question item is used to describe the strength of a question item in revealing the hidden trait being measured. The information value of item number 17 is only 0.10, which is relatively low, with the peak of the curve on the ability axis being too far to the right, this shows that the current information value has not contributed much in revealing the hidden trait being measured.

5) Analysis of item no 26

Question number 26 has a difficulty level of 5.53 on the logit scale. A high level of difficulty indicates that this question is considered very difficult for

students, considering that only 3 out of 250 students answered correctly. This shows that the majority of students have difficulty answering this question, so that item number 26 requires very high student ability in understanding the material being tested. However, when we look at the characteristics of this item based on the Rasch model, it is found that the characteristics do not match the assumptions of the Rasch model.

26	Dewi memberikan hadiah kepada teman di hari ulang tahunnya.Hadiahnya berupa bantal guling yang tebentuk dari gabungan tabung dan kerucut dengan ukuraan tinggi tabung 52 cm, tinggi tabung dan kerucut 62 cm, lebar tabung 14 cm dan Panjang garis miring kerucut 12 , Agar hadiahnya tidak dilihat langsung oleh temannya,Dewi membungkusnya dengan kertas kado.Luas minimal kertas kado yang dibutuhkan Dewi untuk membungkus hadiah tersebut adalah a. 858 cm^2 b. 850 cm^2 b. 94 cm^2 d. 264 cm^2 

Based on analysis using the Rasch model, item number 26 shows a deviation from expectations, indicating the possibility of a problem in the construction of the question. In terms of information function, this question has a low value (0.12), which indicates that the question is less effective in differentiating students based on their abilities. In addition, the difficulty level of this question is too high, but does not provide significant information to measure students' abilities. Therefore, this question needs to be revised or even removed from the test set. If it is maintained, adjustments need to be made to suit the characteristics of the Rasch model.

DISCUSSION

This research shows that the data used is adequate, as shown by the KMO value which exceeds 0.5, and shows that there is a significant relationship between the existing variables. The results of factor analysis identified 11 factors that emerged from the 30 variables studied, with the first factor being the most dominant, marked by an eigenvalue of 4.433. This finding is consistent with previous research findings which also confirmed that the first factor in factor analysis usually has a much greater eigenvalue than the other factors, indicating its dominance in explaining data variations (Ismail & Yoestara, 2022). In addition, the results of factor analysis in this study also support the assumption of unidimensionality, where the first factor makes the largest contribution to understanding the data. This finding is consistent with previous research which also confirmed the importance of the first factor in factor analysis as the dominant factor (Larsen & Warne, 2010).

The visualization in the Scree plot also provides additional evidence about the existence of 11 significant factors in this research instrument. These results contribute to the findings of related research which uses similar methods to identify

the main factors in research instruments (Ayuni & Sari, 2018). Thus, the results of this study provide an important contribution to the understanding of data structure and factors influencing instruments in the context of this study, and are consistent with findings in previous studies.

The findings of this study underline the importance of understanding the level of item difficulty in the analysis of 1PL item response theory and item characteristic curves. From the analysis of thirty questions, most of the questions were in the category of difficulty level which could be considered good, with placement tending to be centered on the medium level of difficulty. However, similar research indicates similar findings, with an emphasis on the importance of balancing the level of difficulty in developing test instruments (Hartono et al., 2022). ICC analysis in this study revealed a pattern that corresponds to the Rasch curve model on items with medium, easy and difficult levels of difficulty. Previous research highlights the important role of balance of difficulty levels in achieving higher measurement validity (Muslihin et al., 2022).

CONCLUSION

1. This research succeeded in showing that the data used was sufficient to carry out factor analysis. This is reinforced by the KMO value which exceeds 0.5, reflecting adequate data quality to reveal the main factors that influence the research instrument. The results of factor analysis highlight the dominance of the first factor, in line with previous research, and the high eigenvalue provides strong evidence for the successful identification and description of these main factors. Understanding this data structure provides a strong basis for continuing the analysis.
2. The importance of understanding the level of item difficulty in the context of item response theory and item characteristic curves. Although most of the items were rated as having a good level of difficulty, this research highlights the need for optimal balance in the development of test instruments. ICC analysis adds confirmation to the pattern fit of the Rasch curve model at various levels of item difficulty, which provides a strong indication of higher measurement quality. These findings consistently support the urgency of achieving a balance of difficulty levels to improve the measurement validity of the instrument.
3. This research makes a significant contribution to the general understanding of data structure and the factors that influence research instruments. The combination of factor analysis and evaluation of item difficulty levels not only validates the instrument, but also provides in-depth insight regarding critical aspects that need to be considered in the development and evaluation of similar instruments in the future.

BIBLIOGRAPHY

Ayuni, N. W. D., & Sari, I. G. A. M. K. K. (2018). Analysis of factors that influencing the interest of Bali State Polytechnic's students in entrepreneurship. *Journal of Physics: Conference Series*, *953*(1). https://doi.org/10.1088/1742-6596/953/1/012071

Djemari. (2017). *Dasar-dasar Evaluasi Pendidikan (edisi revisi)* (2 ed.). Nuha Litera.

Hartono, W., Hadi, S., Rosnawati, R., & Retnawati, H. (2022). Uji Kecocokan Model Parameter Logistik Soal Diagnosa Kemampuan Matematika Dasar. *JNPM (Jurnal Nasional Pendidikan Matematika)*, *6*(1). https://doi.org/10.33603/jnpm.v6i1.5899

Ismail, N. M., & Yoestara, M. (2022). Validity Test Using Principal Component Analysis (PCA) Involving Research Instruments on Learner-Centeredness, Knowledge-Centeredness, and Assessment-Centeredness. *Proceedings of International Conference on Multidiciplinary Research*, *5*(1). https://doi.org/10.32672/pic-mr.v5i1.5255

Javili, A., Dortdivanlioglu, B., Kuhl, E., & Linder, C. (2015). Computational aspects of growth-induced instabilities through eigenvalue analysis. *Computational Mechanics*, *56*(3). https://doi.org/10.1007/s00466-015-1178-6

Kurnia, A. (2019). Analisis Tes Kemampuan Berpikir Kritis Matematika Siswa dengan Menggunakan Generalized Partial Credit Model (GPCM). In *PEDIAMATIKA: Journal of Mathematical Science and Mathematics Education* (Vol. 01, Nomor 02).

Larsen, R., & Warne, R. T. (2010). Estimating confidence intervals for eigenvalues in exploratory factor analysis. *Behavior Research Methods*, *42*(3). https://doi.org/10.3758/BRM.42.3.871

Muslihin, H. Y., Suryana, D., Ahman, Suherman, U., & Dahlan, T. H. (2022). Analysis of the Reliability and Validity of the Self-Determination Questionnaire Using Rasch Model. *International Journal of Instruction*, *15*(2). https://doi.org/10.29333/iji.2022.15212a

Mustafa, P. S., & Masgumelar, N. K. (2022). Kajian Review: Pengembangan Instrumen Penilaian Sikap, Pengetahuan, dan Keterampilan dalam Pendidikan Jasmani dan Olahraga Pinton Setya Mustafa 1 , Ndaru Kukuh Masgumelar 2. *Jurnal ilmiah fakultas keguruan dan ilmu pendidikan*, *8*(1), 31–49.

Nasution, M. Z. (2019). PENERAPAN PRINCIPAL COMPONENT ANALYSIS (PCA) DALAM PENENTUAN FAKTOR DOMINAN YANG MEMPENGARUHI PRESTASI BELAJAR SISWA (Studi Kasus : SMK Raksana 2 Medan). *JURNAL TEKNOLOGI INFORMASI*, *3*(1). https://doi.org/10.36294/jurti.v3i1.686

Pratama, D. (2020). Analisis Kualitas Tes Buatan Guru Melalui Pendekatan Item Response Theory (IRT) Model Rasch. *Tarbawy : Jurnal Pendidikan Islam*, *7*(1). https://doi.org/10.32923/tarbawy.v7i1.1187

Retnawati. (2014). *Teori Respon Butir dan Penerapanya* (3 ed.). Nuhu Medika.

Sanjaya Putra, I. G., & Renda, N. T. (2022). Instrumen Penilaian Sikap Spiritual dan Sikap Sosial Siswa Kelas IV Sekolah Dasar Tema Indahnya Keberagaman di Negeriku. *Jurnal Pedagogi dan Pembelajaran*, *5*(2). https://doi.org/10.23887/jp2.v5i2.46833

Tiara, S. K., & Sari, E. Y. (2019). Analisis Teknik Penilaian Sikap Sosial Siswa Dalam Penerapan Kurikulum 2013 Di Sdn 1 Watulimo. *EduHumaniora | Jurnal Pendidikan Dasar Kampus Cibiru*, *11*(1). https://doi.org/10.17509/eh.v11i1.11905

Zainal, N. F. (2020). Pengukuran, Assessment dan Evaluasi dalam Pembelajaran Matematika. *Laplace : Jurnal Pendidikan Matematika*, *3*(1). https://doi.org/10.31537/laplace.v3i1.310

Zohrabi, M. (2013). Instrument, Validity, Reliability. *Theory and Practice in Language Studies*, *3*(2). https://doi.org/10.4304/tpls.3.2.254-262