Journal of Educational Sciences
Journal homepage: https://jes.ejournal.unri.ac.id/index.php/JES

P-ISSN 2581-1657
E-ISSN 2581-2203

A Modern Approach to the Accuracy of Assessment of Mathematical Creative Thinking Ability of Junior High School Students Through Graded Response Models

Novi Kurniawan, Mujahidawati, Ilham Falani*
Master of Mathematics Education Program, Faculty of Teacher Training and Education, University of Jambi, Jambi City, 36123, Indonesia

ARTICLE INFO

Article history:

Received: 15 May 2025
Revised: 10 June 2025
Accepted: 11 June 2025
Published online: 05 July 2025

Keywords:

Creative Thinking
Mathematics
Valuation
Graded Response Models
Item Response Theory

* *Corresponding author:*

E-mail: ilhamfalani@gmail.com

Article Doi:

https://doi.org/10.31258/jes.9.4.p.2277-2288

A B S T R A C T

The ability to think creatively in mathematics is a crucial aspect that must be possessed by students, so that they can face the dynamics of change and challenges in life that continues to develop. Graded Response Models (GRM) is one of the models of the modern approach of Item Response Theory (IRT) to polytomus data. Graded Response Models (GRM) are used to estimate students' test item ability and mathematical creative thinking skills. This research is a type of quantitative research with an ex post facto method. This study uses the responses of Junior High School (SMP) Class VIII students in the North Bahar sub-district, Muaro Jambi Regency. A total of 120 grade VIII students spread across 4 State Junior High Schools in the sub-district were the subjects of this study. The instrument used was a test instrument for students' mathematical creative thinking skills amounting to 2 questions with each question having 4 indicators of mathematical creative thinking ability. The data was analyzed using the GRM approach using the help of Parscale 4.1 software. Based on the results of data analysis, the results of estimating the ability parameter (θ) of mathematical creative thinking are mostly at medium ability. This can be seen from the histogram of the results of estimating the parameters of mathematical creative thinking ability, most students are in the range of -1 to 1.

1. Introduction

The current era of the industrial revolution 4.0, which is marked by the integration of digital, physical, and biological technologies, requires humans to have new abilities to adapt and compete in the global world. 21st century capabilities are an essential set of capabilities in today's era of technology and information. These abilities include digital literacy, critical thinking, creativity, cooperation, communication, and adaptability (Pare & Sihotang, 2023). One of the abilities that students are indispensable is the ability to think creatively because it is the foundation for various cognitive processes such as designing, solving problems, fixing, and creating new ideas (Halim, A & Ahya Ningsih., 2019).

Thinking ability is an ability that every individual has in finding answers or solutions. One of the thinking skills that is often applied is the ability to think creatively. In the current independent curriculum, the government also promotes the importance of creative thinking skills in learning (Wardani et al, 2023). Creative thinking can be interpreted as producing many relevant ideas/answers, flowing thoughts, generating diverse ideas, different directions of thought, and expanding an idea (Sari & Afriansyah, 2022). To find out the ability to think creatively, there are several indicators that must be met, namely: fluency, flexibility and novelty (Adila et al, 2024).

Currently, the approach in mathematics learning focuses on developing higher-level thinking skills, particularly critical and creative thinking. Creative thinking does not always follow a structured pattern, but rather is a mental habit honed through intuition, stimulation of imagination, exploration of new possibilities, opening up unique perspectives, and the emergence of unexpected ideas. (Pipit Nuur Ainii, 2023). Getting used to connecting different concepts flexibly is a key element in creative thinking. Therefore, it is important to know the extent of students' creative thinking skills in mathematics learning. One method that can be used to measure the level of students' creative thinking is through giving questions which are then analyzed in the assessment process.

Assessment is the process of collecting and processing information to determine the learning needs and developmental achievements or learning outcomes of students (Permendikbudristek. No. 21, 2022). Assessments can be done in various forms, both test and non-test. A test is a procedure for measuring something based on pre-established rules (Nurhasanah et al, 2023). The test instrument used can be in the form of multiple-choice questions or description questions (essay). The administration of this test is expected to be able to produce precise and accurate measurements.

The accuracy of measuring students' creative thinking skills is highly dependent on the quality of the test instruments used. The measurement process should ideally provide an accurate picture of the extent to which students understand the material being studied, so that it can be used as a basis for evaluating the effectiveness of learning and determining the necessary improvement steps (Widiyarto & Inayati, 2023). Each question in the test must go through a rigorous calibration stage to ensure that the questions are of a level of validity and reliability that meets academic measurement standards.

Question calibration is a method to predict the feasibility of a question to measure ability or achievement (Ul Hassan & Miller, 2019). In measurement, there are two approaches that are often used in analyzing question items, namely Classical Test Theory (CTT) and Item Response Theory (IRT).

Classical test theory has long been a reference in the development of measurement tools and instruments. This approach assesses each question item by considering the level of difficulty and differentiation it has. However, the characteristics of the questions are greatly influenced by the individual taking the test, so it can lead to variability in the results obtained. The relationship between the question and the test taker is one of the main weaknesses of classical test theory, because it has the potential to produce less accurate measurements.

To overcome the weaknesses of classical measurement, the use of modern measurements is shown to nullify the inseparability between the group of test participants and the group of test items. Item response theory, hereinafter abbreviated as IRT, does not depend on a specific sample of questions or individuals who take the exam. IRT frees respondents and items from interdependence, so. The difficulty level of the item no longer depends on the ability of the respondent and on the other hand the ability of the respondent no longer depends on the difficulty level of the item (Hambleton, R.K. & Jones, R.W, 1993).

One of the IRT models used is Graded Response Models (GRM). GRM is also known as the Sorted Categorical Response Model because it deals with sorted polytom categories that can relate to constructed response items or selected responses in which test-takers are expected to obtain various score levels such as 0-4 points (Matteucci, M., & Stracqualursi, L, 2006). In this case, the categories are as follows: 0, 1, 2, 3, and 4; and sorted. 'Sorted' means what it says, that there is a certain order or rank of responses.

This study focuses on the accuracy of measuring students' mathematical creative thinking skills by applying the Graded Response Models (GRM) item response theory approach, to analyze polytomy data on essay-shaped questions. The application of this item response theory approach is expected to increase the accuracy of measuring students' mathematical creative thinking skills compared to the use of classical test theory.

2. Methodology

This research is a type of quantitative research with an ex post facto method. Ex post facto research is research that examines causal relationships that are not manipulated or treated by researchers (Nindriyati, 2022). This study uses the responses of Junior High School (SMP) Class VIII students in the even semester of the 2024/2025 school year in the North Bahar sub-district, Muaro Jambi Regency. A total of 120 grade VIII students spread across 4 State Junior High Schools in the sub-district were the subjects of this study. The instrument used was a test instrument for students' mathematical creative thinking abilities with the Pythagorean Theorem material amounting to 2 questions with each question containing 4 indicators of mathematical creative thinking ability. The data analyzed came from students' responses to the mathematical creative thinking ability test in the form of polytomous scores and were analyzed using the GRM approach using the help of Parscale 4.1 software, to obtain an estimate of the parameters of students' mathematical creative thinking ability.

3. Results and Discussion

Results of Model Prerequisite Test Analysis

The analysis of the model prerequisite test includes three main aspects, namely the unidimensional test, the local independence test, and the parameter invariance test, which are interrelated with each other. DeMars (2010) and Saepuzaman et al. (2021) suggest that if the unidimensional assumption is met, then the local independence test will also be automatically met. Meanwhile, according to Dali S. Naga (2012), the unidimensional requirement aims to maintain invariance in the theory of item response. Based on these two statements, it can be concluded that the fulfillment of the one-dimensional test directly ensures the fulfillment of the local independence test and parameter invariance, so that the main focus can be directed to the one-dimensional test only.

The unidimensional test was carried out to ensure that the instrument used could really measure the targeted aspect, namely the students' mathematical creative thinking ability. According to Ilham Falani (2023), unidimensional assumptions can be tested through factor analysis with the help of SPSS software. This test is applied to the data before it is used to estimate the parameters of the test participant's ability. Table 1. Shows the results of the unidimensional test.

Table 1. Unidimensional Test

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.728	59.100	59.100	4.728	59.100	59.100
2	.780	9.748	68.848			
3	.588	7.346	76.195			
4	.529	6.618	82.812			
5	.454	5.672	88.484			
6	.365	4.567	93.051			
7	.297	3.714	96.765			
8	.259	3.235	100.000			

Extraction Method: Principal Component Analysis.

Based on the output of SPSS 27, the results of dimension reduction analysis showed that the principal axis factoring method succeeded in extracting data into several factors that had an eigenvalue greater than one. The data from the analysis obtained the main factors with a total variance explained of 59.1%, while the rest had a contribution of total variance explained which ranged from less than 10%. Here is a plot scree graph for the analysis data. Figure 1. Shows a one-dimensional scree plot visualization.

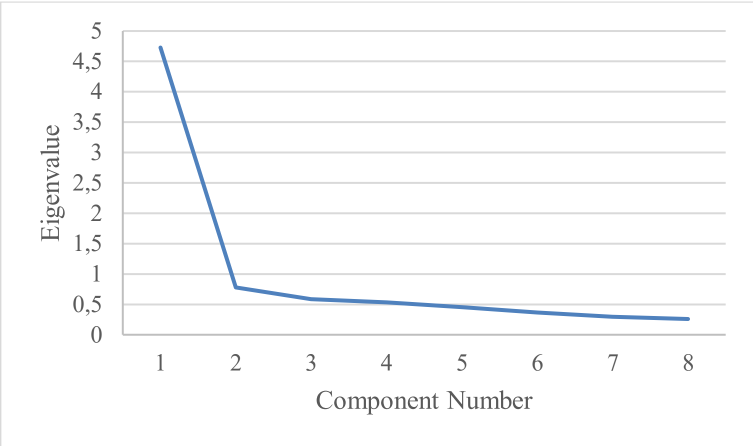

Figure 1. Scree Plot One-Dimensional

Based on the results of factor analysis and Scree Plot visualization in Figure 1, it is known that the main factors in the data have a significant contribution in explaining the total proportion of variance. Therefore, it can be concluded that the test items used have unidimensional characteristics. Most of the question items are combined into a single factor that represents the students' mathematical creative thinking skills.

Model Fit Test Analysis Results

The model fit test is an analysis that aims to evaluate the consistency of individual answers to test questions by referring to the calculation rules that have been set (Falani, 2020). The level of model suitability is analyzed through a prediction evaluation method, which includes examining empirical test results as well as comparison with mathematical expectations of the model (Aji & Retnawati, 2024).

The first step in a fit evaluation is to determine the raw residue, which is the difference between the empirical test results and the mathematical expectations of the model. Then, the raw residue is converted into the raw residue (z). Assuming that the square number of the raw residue follows the khi-squared distribution (χ^2), the fit test is performed using the significance level. This process involves a comparison between the khi-squared value of the residue and the critical value of the khi-squared probability distribution. If the khi-squared value exceeds the critical value, then the question item is considered not to match the model used, while if the value is lower, the item is considered to meet the match. This test item match is performed with the help of Parscale 4.1 software. Table 2. Shows the results of the model fit test.

Based on the output of Parscale 4.1 in Phase 2, an item fit statistics with a GRM χ^2 value of 75.83 (p-value = 0.611) was obtained with a critical distribution χ^2 value of 101.879. This means that the χ^2 value obtained is less than the χ^2 value of its critical distribution, this indicates that the model used has a high match with the empirical data. (Naga, 2012; Retnawati, 2014; Yılmaz, 2019). Furthermore, the model can be used to estimate the parameters of the test participant's ability.

Table 2. Model Fit Test

ITEMS	CHI - SQUARE	D.F.	PROB.
1	6,68274	10.	0,756
2	11,70899	11.	0,386
3	14,12792	12.	0,292
4	13,35268	13.	0,204
5	13,35873	9.	0,146
6	4,94826	9.	0,840
7	4,31848	9.	0,890
8	7,33381	10.	0,695
TOTAL	75,83161	80.	0,611

Results of Estimation of Creative Thinking Skill Parameters

Parameter estimation was carried out on 120 respondents from the results of the creative thinking ability test, data generated from scoring 0, 1, 2, and 3. Parameter estimation is carried out with the help of Parscale 4.1 software located at ph.3.

The normality test was carried out on the results of estimating the parameters of mathematical creative thinking ability with the help of SPSS 27 software. The normality test used is the Kolmogorov-Smirnov normality test. Table 3. Shows the results of the Kolmogorov-Smirnov normality test.

Table 3. Normality Test

			GRM
N			120
Normal Parameters[a,b]	Mean		.000001
	Std. Deviation		1.0041939
Most Extreme Differences	Absolute		.047
	Positive		.038
	Negative		-.047
Test Statistic			.047
Asymp. Sig. (2-tailed)[c]			.200d
Monte Carlo Sig. (2-tailed)[e]	Sig.		.762
	99% Confidence Interval	Lower Bound	.751
		Upper Bound	.773

Based on table 3 above, the analyzed data produced a sig value of > 0.05, which is 0.2. According to the Kolmogorov-Smirnov test criteria, if the sig value > 0.05, then the data is normally distributed (Kadir, 2017). This data can be used for further analysis in the study.

The estimated parameters of mathematical creative thinking ability are said to be low if they are in the range of values between -3 and -1, medium between -1 and 1, and high between 1 and 3. The following is a statistical description of the data results of estimating the parameters of students' mathematical creative thinking ability that has been carried out. Table 4. Show an overview of the distribution of data from the results of estimating the parameters of creative thinking ability analyzed with the GRM model.

Table 4. Results of Estimating Parameters of Creative Thinking Ability with GRM

Persons	**Ability θ**	**Category Ability**	**Persons**	**Ability θ**	**Category Ability**	**Persons**	**Ability θ**	**Category Ability**
1	0,2788	Medium	**41**	-0,0715	Medium	**81**	-0,5341	Medium
2	-0,2908	Medium	**42**	0,3803	Medium	**82**	0,7738	Medium
3	0,5246	Medium	**43**	1,2848	High	**83**	-0,5412	Medium
4	-1,494	Low	**44**	-0,0715	Medium	**84**	1,3948	High
5	-1,1237	Low	**45**	-2,2418	Low	**85**	-1,6678	Low
6	-1,2737	Low	**46**	-1,4707	Low	**86**	0,2933	Medium
7	0,4791	Medium	**47**	-1,1237	Low	**87**	1,6346	High
8	0,7699	Medium	**48**	-0,2908	Medium	**88**	-0,8472	Medium
9	0,5055	Medium	**49**	1,0252	High	**89**	1,6	High
10	-0,8456	Medium	**50**	-0,015	Medium	**90**	-0,7985	Medium
11	-1,1237	Low	**51**	0,3026	Medium	**91**	0,1385	Medium
12	1,2848	High	**52**	-0,076	Medium	**92**	-0,8323	Medium
13	-0,948	Medium	**53**	0,2173	Medium	**93**	0,5261	Medium
14	1,1177	High	**54**	0,4665	Medium	**94**	-0,7736	Medium
15	1,2848	High	**55**	-2,1942	Low	**95**	1,0252	High
16	1,9391	High	**56**	-0,8225	Medium	**96**	-1,1956	Low
17	-0,3056	Medium	**57**	0,867	Medium	**97**	-0,5341	Medium
18	-0,5362	Medium	**58**	0,7352	Medium	**98**	-0,4672	Medium
19	-0,8682	Medium	**59**	-1,9056	Low	**99**	-0,2522	Medium
20	-1,1046	Low	**60**	-0,6677	Medium	**100**	-1,405	Low
21	-0,6013	Medium	**61**	-0,6879	Medium	**101**	1,6086	High
22	0,3006	Medium	**62**	-0,0204	Medium	**102**	-0,3126	Medium
23	2,0698	High	**63**	-0,0908	Medium	**103**	1,2484	High
24	0,1409	Medium	**64**	-0,8125	Medium	**104**	0,6549	Medium
25	-1,9269	Low	**65**	0,6343	Medium	**105**	0,4441	Medium
26	0,2369	Medium	**66**	0,0307	Medium	**106**	1,14	High
27	0,9647	Medium	**67**	-0,0065	Medium	**107**	1,8656	High
28	0,1409	Medium	**68**	0,3585	Medium	**108**	0,0908	Medium

29	-2,1942	Low	**69**	-1,9269	Low	**109**	-0,688	Medium
30	-1,4699	Low	**70**	0,7352	Medium	**110**	-0,6097	Medium
31	-0,1843	Medium	**71**	-0,9822	Medium	**111**	1,0614	High
32	-0,8415	Medium	**72**	1,0191	High	**112**	1,4277	High
33	-0,925	Medium	**73**	-0,2262	Medium	**113**	-1,592	Low
34	1,0614	High	**74**	0,5462	Medium	**114**	0,3399	Medium
35	-0,4753	Medium	**75**	0,9222	Medium	**115**	0,4847	Medium
36	0,9863	Medium	**76**	0,3898	Medium	**116**	0,8583	Medium
37	0,5261	Medium	**77**	1,8171	High	**117**	0,0967	Medium
38	-1,656	Low	**78**	1,4482	High	**118**	-0,154	Medium
39	-0,1922	Medium	**79**	0,1894	Medium	**119**	1,3362	High
40	0,8106	Medium	**80**	0,2933	Medium	**120**	0,1613	Medium

Figure 2. Showing the visualization of the distribution of the data can also be seen from the histogram

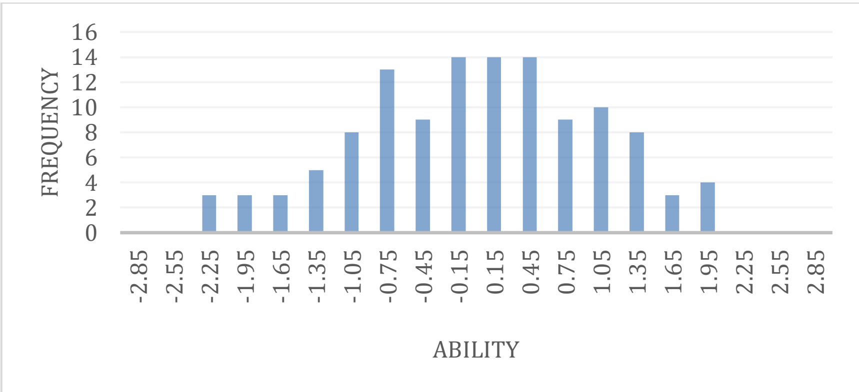

Figure 2. Histogram Results of Estimation of Parameters of Mathematical Creative Thinking Ability

From the results of estimating the parameters of mathematical creative thinking ability parameters of 120 students, 19 students with low ability, 79 medium ability students, and high ability 22 students were obtained.

Results of Parameter Estimation of Question Item

Parameter estimation of a question item is a parameter that describes the differential power (a) and the level of difficulty of a question item (b). The results of the analysis are found in the output of Parscale 4.1 in Ph.2, Table 5. Shows the parameters of the mathematical creative thinking ability test item. Differential power (a) that is considered good has a value range between 0.25 to 1.5 as a typical range, while a value above 1.5 is also included in the good category. On the other hand, values below 0.25 are categorized as low (Ridlo, 2012; Hambleton et al., 1991).

Table 5. Parameters of Mathematical Creative Thinking Ability Test Items

Grain	Differentiation (a)	Difficulty Level (b)
1	1,887	0,773
2	1,368	0,197
3	1,207	1,064
4	1,292	-0,438
5	1,119	-0,558
6	1,476	-0,667
7	0,958	-0,365
8	1,019	-0,253

Based on table 5, it can be seen that no item has a difference power (a) below 0.25, so that the parameters of this item are declared to meet good standards. As for the difficulty level (b), the value is considered good if it is in the logit range of -3 to 3 (Hambleton et al., 1991; Retnawati, 2014; Saepuzaman et al., 2021). From table 5, it can be seen that the difficulty level in this instrument is relatively good, since all the logit values of the item are in the range of -3 to 3.

Test *Information Function*

The function of test information plays a role in describing the characteristics of students' abilities based on the measurement results (Supriyati et al, 2021). The curve of the test information function can be divided into four main patterns: (1) if the curve is tilted to the left, it means that the majority of students being tested have low ability, (2) if the curve is tilted to the right, then most of the students have high ability, (3) if the curve forming a peak in the middle resembles a normal distribution, then the ability of the students being measured tends to be at a moderate level, and (4) if the curve decreases in the middle and rises on both sides, indicating the presence of two groups of students with high and low ability.

In addition, the item information function has an inversely proportional relationship to the standard error measurement. The smaller the error in the measurement, the greater the information obtained from the test item (Falani, 2020). Figure 3. Indicates the function of test item information.

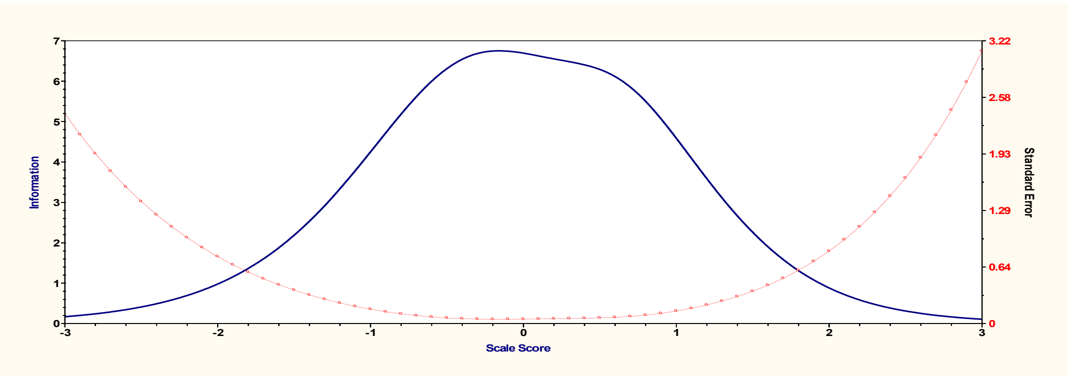

Figure 3. Test Information Function Graph

From Figure 3 above, this graph is mountainous in the middle with a peak that is in the range of values around -1 to 1 and the information provided by the test questions reaches a peak of around 6.8. This means that the test questions given are most accurate when used for students with a moderate level of creative thinking ability. That way, students who have creative thinking skills in that range, the test results will be more precise and trustworthy.

The graph also shows that outside the peak range, either towards lower (less than -1) and higher (above 1) values, the amount of information generated by the test has decreased significantly. This shows that for students with creative thinking skills that are too low or too high, the test questions given are less effective in producing accurate information. Therefore, the mathematical creative thinking skills test that was compiled is more suitable to assess students with a level of creative thinking ability that is in the medium category.

4. Conclusion

This research is a quantitative research with an ex-post facto method. This study aims to improve the accuracy of students' assessment of mathematical creative thinking skills by using the modern IRT Graded Response Models approach. The instrument consisting of 2 essay questions with 4 indicators of creative thinking ability each met the assumption test, so it is suitable for analysis with item response theory. The results of the study showed that the model match on the test device used Graded Response Models (GRM). All questions have met the GRM standards and are declared fit so that they meet the empirical validity that can measure students' mathematical creative thinking skills, the test instrument also has a difference and a good level of difficulty. The test information function describes the mathematical creative thinking ability test that was compiled is more suitable for assessing students with the level of creative thinking ability that is in the medium category. The results of the estimation of the ability parameter (θ) to think creatively mathematically, most students are at moderate ability. Therefore, a modern approach to GRM item response theory is recommended for estimating students' creative thinking skills on tests with polytomus scores. The results of the right estimation greatly affect the assessment carried out by the teacher, Assessment that is carried out correctly can help teachers in evaluating the learning process and improving the quality of their teaching. For further research, it is recommended to examine the estimation of students' abilities with the theory of item response with other models.

References

Adila, H. I., Iriani, D., & Falani, I. (2024). Analisis Kemampuan Berpikir Kreatif dalam Pemecahan Masalah Matematika Berdasarkan Tipe Kepribadian Ekstrovert dan Introvert Siswa. *JURNAL PENDIDIKAN MIPA*, *14*(1), 200–211.

Aji, N. W., & Retnawati, H. (2024). Perbandingan Kecocokan Model Analisis Teori Respons Butir Pada Data Politomi. *Jurnal Statistika dan Sains Data*, *1*(2), 31-47.

DeMars, C. (2010). *Item response theory*. Oxford University Press.

Falani, I. (2023). Desain dan Validasi Aplikasi Tes Literasi Matematika Berbasis Komputer dengan Pendekatan Item Response Theory: Design and Validation of Computer-Based Test of Mathematical Literacy with Item Response Theory Approach. *Edu-Sains: Jurnal Pendidikan Matematika dan Ilmu Pengetahuan Alam*, *12*(2), 16-28.

Falani, I., Akbar, M., & Naga, D. S. (2020). The Precision of Students' Ability Estimation on Combinations of Item Response Theory Models. *International Journal of Instruction*, *13*(4), 545-558.

Halim, A., & Ahyaningsih, F. (2019). Pengaruh Pendekatan Matematika Realistik Terhadap Kemampuan Berfikir Kreatif Dan Pemecahan Masalah Matematika Siswa Kelas VII. *Jurnal Paradikma*, *12*(02), 1-9.

Hambleton, R.K. & Jones, R.W. (1993). Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development. ITEMS – Instructional Topics in Educational Measurement, 12(3), 38 – 47. https://doi.org/10.1111/j.1745 – 3992.1993.tb00543.x

Hambleton, R. K., Swaminathan, H., & Rogers, D. J. (1991). *Fundamentals of Item Response Theory*. Sage Publications.

Kadir. (2017). *Statistika Terapan*. (5, Ed.) (3rd ed.). Jakarta: Rajawali Press.

Matteucci, M., & Stracqualursi, L. (2006). Student assessment via graded response model. *Statistica*, *66*(4), 435-447.

Naga, D. S. (2012). *Teori Sekor pada Pengukuran Mental*. Jakarta: PT Nagarani Citrayasa.

Nindriyati, D. (2022). Hubungan kecerdasan logis matematis dengan hasil belajar matematika. *Instruksional*, *3*(2), 187-196.

Nurhasanah, N., Azhari, A., Berutu, K., Putra, T. J., Hasibuan, R. H., & Nasution, I. (2023). Evaluasi Pembelajaran dikelas. Jurnal Motivasi Pendidikan Dan Bahasa, 1(2), 257–270.

Pare, A., & Sihotang, H. (2023). Pendidikan Holistik untuk Mengembangkan Keterampilan Abad 21 dalam Menghadapi Tantangan Era Digital. *Jurnal Pendidikan Tambusai*, *7*(3), 27778–27787. https://doi.org/10.31004/jptam.v7i3.11268

Pipit Nuur Ainii, P. I. P. I. T. (2023). Pengaruh Model Pembelajaran Inkuiri Sosial terhadap Kemampuan Berpikir Kreatif Siswa di Sekolah Menengah Atas Negeri 3 Tapung (Doctoral dissertation, Universitas Islam Negeri Sultan Syarif Kasim Riau).

Retnawati, H. (2014). Teori respons butir dan penerapannya: Untuk peneliti, praktisi pengukuran dan pengujian, mahasiswa pascasarjana. *Yogyakarta: Nuha Medika*.

Ridlo, S. (2012). Pengembangan Tes Pengetahuan Praktikum Biologi Berdasarkan Graded Response dan Generalized Partial Credit. *Jurnal Penelitian dan Evaluasi Pendidikan*, *16*, 166–182.

Saepuzaman, D., Istiyono, E., Haryanto, H., Retnawati, H., & Yustiandi, Y. (2021). Analisis karakteristik butir soal fisika dengan pendekatan IRT penskoran dikotomus dan politomus. *Radiasi: Jurnal Berkala Pendidikan Fisika*, *14*(2), 62–75. https://doi.org/https://doi.org/10.37729/radiasi.v14i2.1200

Sari, R. F., & Afriansyah, E. A. (2022). Kemampuan berpikir kreatif matematis dan belief siswa pada materi persamaan dan pertidaksamaan linear. *Plusminus: Jurnal Pendidikan Matematika*, *2*(2), 275-288.

Supriyati, Y., Falani, I., & Maulana, S. (2021, March). The information function of mixed-format test of physics learning outcomes measurement. In *AIP Conference Proceedings* (Vol. 2320, No. 1). AIP Publishing.

Ul Hassan, M., & Miller, F. (2019). Optimal Item Calibration For Computerized Achievement Tests. Psychometrika, 84(4), 1101-1128.

van der Linden, W. J., & Hambleton, R. K. (2013). Handbook of modern item response theory. Springer Science & Business Media.

Wardani, S., Asbari, M., & Misri, K. I. (2023). Pendidikan yang Memerdekakan, Memanusiakan dan Berpihak pada Murid. *Journal of Information Systems and Management (JISMA)*, *2*(5), 35-43.

Widiyarto, A., & Inayati, N. L. (2023). Penerapan Evaluasi Pembelajaran Tes Dan Non-Tes Pendidikan Agama Islam di Sekolah Menengah Kejuruan. *Munaddhomah: Jurnal Manajemen Pendidikan Islam*, *4*(2), 307-316.

Yılmaz, H. B. (2019). A Comparison of IRT Model Combinations for Assessing Fit in a Mixed Format Elementary School Science Test. *International Electronic Journal of Elementary Education*, *11*(5), 539–545. https://doi.org/10.26822/iejee.2019553350

How to cite this article:

Kurniawan, N., Mujahidawati., & Falani, I. (2025). A Modern Approach to the Accuracy of Assessment of Mathematical Creative Thinking Ability of Junior High School Students Through Graded Response Models. *Journal of Educational Sciences*, *9(4)*, 2277-2288.