Proceeding Konferensi Nasional Matematika (KNM) XXII
Universitas Bengkulu, 15-16 Juli 2024
Dipublikasikan online pada:
PATTIMURA PROCEEDING: Conference of Science and Technology
e-ISSN: 2829-3770, volume 5 issue 1, pp.41-48, Edisi November 2024

Validity of the Rasch Model-Based Mathematical Critical Thinking Ability Instrument

Dinda Aulia Mawaddah[1], Ayu Faradillah[2*], Fitri Alyani[3]

[1,2,3]Faculty of Teacher Training and Education, Universitas Muhammadiyah Prof. DR. HAMKA Jalan Tanah Merdeka No. 20, RT.11/RW.2, Rambutan, Ciracas Sub-district, East Jakarta City, Special Capital Region of Jakarta 13830, Indonesia

Corresponding author's e-mail: **** ayufaradillah@uhamka.ac.id***

Keywords:
Validity;
Rasch model:
Critical Thinking.

ABSTRACT

Critical intellectual skills are very necessary to make a decision, but students' critical intellectual abilities are still very low due to students' lack of ability to solve problems. This research aims to determine the results of instrument validation of applications of the Rasch model to critical thinking abilities. This study employs survey methodologies together with a quantitative methodology. The subjects of this research consisted of 134 high school students consisting of class XI. The results of this study showed two valid items, with one misfit item and five misfit items, while the DIF test showed all items indicated DIF. Testing the reliability of the instrument with the Winsteps application, obtained reliability with a Cronbach Alpha (KR-20) of 0.87, it appears that 0.87 is greater than 0.7 thisfalls into the category of excellent. People's 0.84 dependability score falls into the excellent range. An item's dependability of 0.98 falls within the excellent range. The 2.31person separation falls within the "good" category.

DOI: https://doi.org/10.30598/ppcst.2024.knmxxii.41-48

1. INTRODUCTION

Students need to be able to think critically about mathematics in order to be able to make decisions and behave logically, namely the need for strategies and also consideration in solving a problem in making answers that are carried out logically, critically and systematically [1][2]. Apart from that, the ability to think critically mathematically or called high-order mathematical thinking skill, is an ability that must be possessed by each individual student as a goal to solve mathematical problems in everyday life [3][4].

Developing critical thinking ability instruments in mathematics is very important because it is a high-level thinking skill that needs to be instilled in students. Meanwhile, this ability helps them in solving problems, reasoning, communicating, making connections, and learning in a more holistic and organized manner [5], [6]. However, in research conducted by Septiana, the results of research on critical thinking skills in mathematics in junior high school students in Bandung on flat-sided building material, with the average percentage value of all indicators being below 50%, it means that critical thinking skills are still very low. Because when learning, students really need something that supports improving student competence. Thus, learning outcomes can be better, especially in critical mathematical thinking skills [4].

Accurately measuring mathematical critical thinking ability instruments requires valid but also reliable assessments. Validity is a measuring tool (instrument) used in research to measure what is to be measured that can be accepted according to standards. Validity tests can be measured through person correlation or product moment correlation. If the instrument data is measured, the validity of the instrument is high, it will show accuracy [7][8]. Validity testing is important to determine the suitability of the instrument used so as not to produce distorted data, meaning the instrument must be valid with the measuring instrument used and the data must be valid [7][9]. Furthermore, Rasch modeling is one of the most well-known models in Item Response Theory (IRT). The basis of the Rasch model is a probabilistic concept which states that individuals with a higher level of ability than others ought to be more likely to respond to the query correctly. Likewise, more difficult questions will reduce an individual's chances of answering them correctly. According to Sumintono and Widhiarso, the Rasch model is a very useful analytical tool for testing the validity and reliability of instruments, as well as assessing persons and items simultaneously. The Rasch model fulfills five measurement principles, namely: (a) is able to provide linear measurements with consistent intervals, (b) can overcome missing data, (c) can provide more accurate estimates, (d) is able to detect model inaccuracies, and (e) offer measuring tools that are unaffected by the parameters under study [10].

Based on previous research which is in line with each other, including research on the Development of Authentic Assessment Instruments for Critical Thinking Skills Tests in Mathematics Learning. Empirical validity is proven and analyzed using the Winsteps 3.73 program. The polytomous data from the trial results were analyzed using the Rasch model with the help of the Winsteps 3.73 program. The results of the research show that the amount of Person Reliability in Mathematics students' critical thinking skills is 0.81 while Item Reliability is 0.91. The magnitude of Cronbach's Alpha is 0.82. Students' average critical thinking capacity on the items is indicated by the Person Measure 0.16 logit. Or in other words, students tend to answer more correctly than all the questions they do [11]. The next research is Mathematical Resilience: Validity and Reliability with Rasch Model and SPSS in Senior High School. This research aims to examine the validity of a mathematics questionnaire using the Winsteps 3.73 application. The results showed that there were three items that were invalid, and six items that only met one validity requirement. Therefore, of the total existing items, 26 items were retained because they met the criteria. With the Winsteps application, the mathematics questionnaire was declared valid and reliable [12]. The instrument validation sheet for critical and creative thinking abilities was assessed using Rasch via Facets software. The analysis results show that this instrument is valid because it is fit. The reliability of the instrument is in the very good group, with the critical thinking ability instrument's Cronbach's Alpha value of 0.87 and the creative thinking ability instrument's score of 0.95. [13]. Furthermore, research on the Effect of Validity, Reliability and Level of Difficulty on the Quality of Economic Question Items Using Anates Software methods for gathering data that were employed in this study include observation, documentation and tests. The study's findings indicated that out of all the questions examined, 17 questions (57%) were declared valid, while 13 questions (43%) were invalid. Reliability shows a high category with a value of 0.73, which is greater than 0, meeting the required category [14].

There hasn't been much study on validity and reliability assessments to gauge students' mathematical critical thinking skills using the Rasch Model and Winsteps, according to a number of conducted studies.to find out whether students' critical thinking abilities are high or low [15]. Therefore, researchers want to know the critical thinking skills in mathematics of high school pupils high by measuring the validity and reliability of instruments using the Rasch Model. Next, the problem formulation of this research is how to validate mathematical critical thinking questions using the Rasc model.

2. RESEARCH METHODS

2.1 Research Design Participants

This research uses a quantitative approach carried out systematically and continuously, data in the form of exact numbers, data collection using instruments, and emphasis on data analysis [16]. Namely by asking questions in the form of essays to all grade 11 student respondents, so that this research possesses data comparable to all current samples. The sensory survey method aims to see the situation that is the object of research [17]. The survey the researchers used is a valid survey method and is also modified according to the research objectives and also in accordance with the data collected.

According to Sugiyono, Purposive sampling is a method by which researchers select and collect samples based on specific criteria [18]. Depending on the demands of the research to be conducted, many types of research can be conducted utilizing this purposive sampling technique. Specifically, 134 students made up the research sample.

Table 1. Data on the number of students based on class, age, place of residence and gender

Demographics	Description	Code	Amount
Class	XI	A	134
Age	16-17	B	120
	18-19	C	14
Domicile	Jakarta	D	112
	Depok	E	23
	Bogor	F	1
Gender	Female	G	85
	Male	H	49

Table 1 shows descriptions, codes and quantities to determine demographic data consisting of data groupings, such as one class, two ages, three domiciles, and two gender.

2.2 Instruments

The tools used in this study are based on measures of critical thinking in mathematics. With a test instrument of three questions, descriptions of the material on systems of linear equations with three variables. In this indicator there are four levels of mathematical critical thinking abilities including inter-hacking, analysis, evaluation, inference.

Table 2. Indicators Consist of Interpretation, Analysis, Evaluation, Inference

Indicators	Information	Score
Interpretation	Not conveying information that is already known and asked about	0
	Convey information that is already known and asked appropriately	1
	Convey information that is known correctly or information that is asked correctly.	2
	Convey information known from the question accurately but not completely.	3
	Convey information that is known and asked from questions completely and accurately.	4
Analysis	Do not formulate a mathematical model from the questions provided.	0
	Formulate a mathematical model from the questions provided but it is not accurate.	1
	Formulate mathematical models from the questions provided accurately without explanation.	2

Indicators	Information	Score
	Formulate the mathematical model of the problem provided accurately but the explanation is incorrect.	3
	Formulate mathematical models of the questions provided accurately and provide correct and complete explanations.	4
Evaluation	Does not apply strategies in solving problems.	0
	Applying inappropriate or incomplete strategies to solve problems.	1
	Applying the right strategy to solve the problem but incomplete, or using an inappropriate but complete strategy.	2
	Applying the right strategy to solve the problem completely but making mistakes in calculations or explanations.	3
	Applying the right strategy in solving problems completely and correctly in carrying out calculations or explanations.	4
Inference	Didn't conclude.	0
	Drawing wrong conclusions that are not appropriate to the context.	1
	The conclusion is incorrect even though it is related to the context.	2
	The conclusion is appropriate to the context but not complete.	3
	Conclude accurately and completely according to the context of the question.	4

Table 2 describes the indicators and assessment guidelines for the mathematical creative thinking ability test instrument used in assessing student work results. This instrument has gone through a content validation process, It evaluates how closely the exam captures students' comprehension of the subject matter in relation to the curriculum. An academic and a math instructor were the two specialists that conducted the validation process. These are the outcomes of this validation.

Mrs. Ani wants to buy fruit parcels for a charity event at her house. The first parcel contains 2 kg of apples, 1 kg of oranges and 3 kg of guavas at a price of Rp. 106,000, then the second parcel contains 2 kg of guavas and 2 kg of oranges at a price of Rp. 64,000 because Mrs. Ani knew that many people would come, Mrs. Ani bought additional fruit, 3 kg of apples, 2 kg of oranges for IDR. 90,000. Determine the price per kilogram for each fruit	Mrs. Ani wants to buy fruit parcels for a charity event at her house. The first parcel contains 2000 grams of apples, 1000 grams of oranges and 3 kg of guavas at a price of IDR. 106,000, then the second parcel contains 2000 grams of guava and 2 kg of oranges at a price of Rp. 64,000 because Mrs. Ani knew that many people would come, Mrs. Ani bought additional fruit, 3000 grams of apples, 2000 grams of oranges for Rp. 90,000. Determine the price per kilogram for each fruit
(a)	**(b)**

Figure 1. (a) before validation (b) after validation

In Figure 1.(a) before validation explains about Mrs. Ani buying fruit for a celebration event, the amount of fruit is a lot of fruit, the load value is in kilograms. Then Figure 1. (b) explains the results after validation, namely the number of loads of fruit, not just kilograms, but there must be different variations, namely in this question using grams and kilograms

Validation sheet for test grid instruments, test questions and answer criteria for experts. In this validation stage, experts assess 3 questions with each item having 12 aspects of the critical thinking ability instrument. Each aspect can receive a minimum of 1 and a maximum of 4,and zero is no value. A response with a score of 4 indicates that the thinks the question is appropriate to the topic with improvements. The answer with a score of 0 shows no response thinks the question is not appropriate to the topic and needs to be corrected. Experts provide opinions that the prototype questions can be used without revision, there are some components of the questions that need to be revised, or all components of the questions need to be revised.

3. RESULTS AND DISCUSSION

3.1 Validity

The Rasch Model can be used to assess the validity and reliability of the mathematical critical thinking ability exam. Furthermore, this research to measure validity uses the Rasch model with the Winsteps application. An item is said to be valid because the item has the validity criteria applied in Rasch analysis

3.1.1 Item Fit

The manner in which to which each instrument item fits with the chosen measurement model is referred to as item fit. Every test question is guaranteed to measure construct validity through item appropriateness analysis.[19]. Three primary factors are used to determine an item's suitability: the Outfit mean-square value (MNSQ), the Outfit z-standard value (ZSTD), and the Point Measure Correlation value (PT-MEASURE CORR). An item person is said to be fit if it meets the criteria, namely Outfit MNSQ has a value of 0.5-1.5, Outfit ZSTD has a value of -2.0-2.0, and Pt Mean Corr has a value of 0.4-0.85.[12] This Table presents the findings of validity and reliability tests that were examined using the Rasch Model.

Table 3. Item Validity Analysis Results

Entry Number	Item	MNSQ	ZSTD	Pt Mean Corr
1	P1	1.95	5.7	0.34
2	P2	1.01	0.1	0.64
3	P3	1.05	0.4	0.63
4	P4	1.05	0.3	0.55
5	P5	1.06	0.4	0.58
6	P6	0.93	-0.5	0.70
7	P7	0.75	-2.1	0.74
8	P8	0.75	-1.4	0.61
9	P9	0.84	-1.1	0.65
10	P10	0.93	-0.5	0.71
11	P11	0.79	-1.7	0.74
12	P12	0.75	-1.7	0.67

In Table 3 there are twelve assessments, there is one assessment that is not fit, namely question number 1, because it does not meet the three fit item criteria, while in item 7 the ZSTD outfit value is outside the -2.0-2.0 criteria, so it can be concluded that the question items used are valid.

3.1.2 Person Fit

Test-taker behavior can be managed by employing the person-fit approach. This is significant since the test seeks to gauge test takers' comprehension of the subject matter, not just their ability to provide accurate responses. In addition to making objective conclusions, this approach helps to maintain measurement validity during test implementation [20].

Table 4. Misfit Order Person Fit Output Results on Winsteps

No	Code person	MNSQ	ZSTD	Pt Mean Corr
1	071ACGJ	**6.66**	**6.5**	**-0.16**
2	001ACHI	**2.85**	**2.2**	**0.18**
3	013ADHJ	**3.15**	**3.7**	**0.04**
4	023ACHJ	**2.53**	**3.0**	**-0.14**
5	034ACGI	**2.43**	**2.1**	**0.09**
6	130ABHJ	**1.85**	1.8	**0.31**
7	056ACHJ	**1.74**	1.4	**0.20**
8	055ACHJ	**1.74**	1.6	**0.37**
9	063ACHJ	**1.63**	0.9	**0.15**
10	019ACGJ	**1.89**	1.7	0.44
11	066ACHI	**1.74**	1.6	0.42
12	106ADHJ	**1.91**	1.7	**0.31**
13	016ACHJ	**1.85**	2.0	0.47
14	076ACGJ	**1.55**	1.2	0.43

No	Code person	MNSQ	ZSTD	Pt Mean Corr
15	119ACHJ	1.36	0.8	**0.34**
16	009ACGJ	**1.53**	1.2	**0.03**
17	048ACHI	**1.51**	1.2	**0.01**
18	133ACHJ	1.26	0.7	**0.36**
19	104ACHI	1.29	0.6	**0.25**
20	070ACHJ	1.41	1.0	**0.23**
21	029ACHI	1.44	1.2	**0.35**
22	087ABGJ	1.29	0.8	**0.30**
23	085ACGJ	1.28	0.8	**0.33**
24	005ACHI	0.80	-0.1	**0.16**
25	123ACHJ	0.78	-0.3	**0.31**
26	084ACHI	0.65	-0.4	**0.31**
27	047ADHI	0.66	-0.3	**0.30**
28	041ADHI	0.64	-0.6	**0.21**
29	030ACHJ	0.54	-0.8	**0.33**
30	093ACHI	0.51	-1.0	**0.37**
31	011ACHJ	**0.46**	-1.4	0.67
32	035ACHJ	**0.46**	-1.0	0.48
33	075ACHI	**0.46**	-1.7	0.79
34	080AEHJ	**0.44**	-1.2	0.64
35	050ACHI	**0.44**	-1.7	0.80
36	124ACGJ	**0.44**	-1.5	0.52
37	122ACHJ	**0.43**	-1.3	0.69
38	031ACHI	**0.43**	-1.8	0.80
39	064ADHI	**0.42**	-1.9	0.60
40	111ACHJ	**0.43**	-1.5	0.52
41	073ACHJ	**0.41**	-1.7	0.83
42	113ACHJ	**0.40**	-2.0	0.65
43	008ACHI	**0.37**	-1.3	0.54
44	025ACHI	**0.36**	-1.3	0.63
45	042ADHJ	**0.34**	-1.4	0.65
46	120ACGJ	**0.30**	**-2.2**	0.78
47	115ABHJ	**0.29**	-1.6	0.72

Table 4 shows the person fit results. Shows person fit results that satisfy the requirements for Point Measure Correlation (PTMEA-CORR), Outfit Z-Standardized Values (ZSTD), and Outfit Mean Square (Outfit MNSQ). Namely, in the person code (1, 13, 23, 34, 71) there are five that are invalid. Additionally, six person codes (9, 48, 55, 56, 63, 106) that indicated misfit were found in the statement results for those who did not meet the ZSTD criteria but did meet the MNSQ and PTMEA-CORR criteria. Meanwhile, those who do not meet the PTMEA-CORR criteria but meet the MNSQ and ZSTD are person code at 120 which means misfit. Then only one of the criteria from PTMEA-CORR is found in the person codes (5, 29, 30, 41, 62, 70, 85, 87, 113, 119, 104, and 123) then the rest only meet the MNSQ criteria [12].

3.2 Reliability

Reliability test research using the reliability analysis output display implementing the Winsteps software with the Rasch model. By using a summary table, namely in the Cronbach Alpha (KR-20) section, item reliability, person reliability, person separation, and item departure. As follows below is the reliability analysis output Table [12].

Table 5. Review of Output Summary Statistics

Statistics	Value
Alpha Cronbach (Kr-20)	0.87
Person Reliability	0.84
Item Reliability	0.98
Person Separation	2.31
Item Separation	6.52

Based on Table 5, the Cronbach Alpha (KR-20) instrument value in the research is 0.87, it can be seen that 0.87 is greater than 0.7, which is in the very good category. Person reliability falls into the good category with a 0.84. With an item dependability of 0.98, it is considered very good. A person separation of 2.31 is considered to be in the good range. Item separation was 6.53 in the very good category [12].

3.3 Differential Item Function (DIF)

Validation of the Differential Item Function (DIF) content used in this research. In Rasch modeling, DIF [20]. used to identify items in an instrument that demonstrate bias or provide an advantage to certain demographic groups. To determine the presence of DIF, an item must have a probability of less than 0.05 and a DIF contrast value larger than 0.5 [22].

Table 6. Differential Item Function

No	Item	DIF Measure	DIF Contrast	t	Probability
1	P1	-0.30	-0.09	-0.41	0.9866
		-0.39	0.09	0.41	
2	P2	-0.56	0.40	1.90	0.0544
		-0.16	-0.40	-1.90	
3	P3	-0.33	-0.04	0.08	0.9669
		-0.29	-0.04	-0.08	
4	P4	1.61	-0.03	-0.08	0.9143
		1.58	0.03	0.08	
5	P5	0.15	-0.17	-0.71	0.4660
		-0.01	0.17	0.71	
6	P6	-0.66	0.03	0.14	0.2869
		-0.63	-0.03	-0.14	
7	P7	-0.63	-0.17	-0.85	0.9487
		-0.80	0.17	0.85	
8	P8	1.25	-0.10	-0.32	0.6817
		1.15	0.10	0.32	
9	P9	0.03	0.06	0.25	0.8879
		0.09	-0.06	-0.25	
10	P10	-0.50	0.05	0.24	0.7050
		-0.45	-0.05	-0.24	
11	P11	-0.66	-0.09	-0.44	0.2858
		-0.75	0.09	0.44	
12	P12	0.66	-0.05	-0.18	0.2371
		0.62	0.05	0.18	

Based on the data in Table 6, of the twelve items to fulfill the Differential Item Function (DIF) which are in accordance with the criteria in the DIF Contrast item, none of them meet the criteria because the results are less than 0.5, then for Probability neither does not meet the criteria because it is more than 0.05 This means that the conclusion obtained, if all items are unbiased, means the statement is valid.

4. CONCLUSIONS

There were twelve assessments in the fit items using the Rasch Model in Winsteps, according to the computation of the validity of the mathematical critical thinking capability test instrument using the Rasch Model, there was one assessment that was not fit, namely question number 1, because it did not meet the three fit item criteria, whereas in item 7 the ZSTD outfit value is outside the -2.0-2.0 criteria, so it can be concluded that the question items used are valid. whereas in the Differential Item Function (DIF) of the twelve items to fulfill the DIF which corresponds to the criteria in the DIF Contrast item, none of them met the criteria because the results were less than 0.5, then for Probability neither did it meet the criteria because it was more than 0.05. This means that the conclusion obtained, if all items are unbiased, means the statement is valid.

REFERENCES

[1] Melyana and H. Pujiastuti, "Pengaruh Kepercayaan Diri Terhadap Kemampuan Berpikir Kritis Matematis Siswa Smp," JPMI (Jurnal Pembelajaran Mat. Inov., vol. 3, no. 3, p. 244, 2020, doi: 10.22460/jpmi.v3i3.239-246.

[2] R. Septiana, Y. S. Febriarini, and L. S. Zanthy, "Analisis Kemampuan Berpikir Kritis Matematis Siswa SMP," J. Pembelajaran Mat. Inov., vol. 2(6), no. 6, pp. 393–399, 2019.

[3] D. Ratnawati, I. Handayani, and W. Hadi, "Pengaruh Model Pembelajaran Pbl Berbantu Question Card Terhadap Kemampuan Berpikir Kritis Matematis Siswa Smp," Edumatica J. Pendidik. Mat., vol. 10, no. 01, pp. 44–51, 2020, doi: 10.22437/edumatica.v10i01.7683.

[4] A. Yanwar and A. Fadila, "Analisis Kemampuan Berpikir Kritis Matematis : Dampak Pendekatan Saintifik ditinjau dari Kemandirian Belajar," Desimal J. Mat., vol. 2, no. 1, pp. 9–22, 2019, doi: 10.24042/djm.v2i1.3204.

[5] F. Az Zahra and D. L. Hakim, "Kemampuan Berpikir Kritis Matematis Siswa Sma Pada Materi Bangun Ruang Sisi Datar Pasca Pembelajaran Jarak Jauh," Teorema Teor. dan Ris. Mat., vol. 7, no. 2, p. 425, 2022, doi: 10.25157/teorema.v7i2.7221.

[6] Y. Risah, Sutirna, and D. L. Hakim, "Pencapaian Kemampuan Berpikir Kritis Matematis Siswa Pada Materi Trigonometri," JPMI (Jurnal Pembelajaran Mat. Inov., vol. 4, no. 2, p. 308, 2021, doi: 10.22460/jpmi.v4i2.307-316.

[7] L. Amanda, F. Yanuar, and D. Devianto, "Uji Validitas dan Reliabilitas Tingkat Partisipasi Politik Masyarakat Kota Padang," J. Mat. UNAND, vol. 8, no. 1, p. 179, 2019, doi: 10.25077/jmu.8.1.179-188.2019.

[8] H. Puspasari and W. Puspita, "Uji Validitas dan Reliabilitas Instrumen Penelitian Tingkat Pengetahuan dan Sikap Mahasiswa terhadap Pemilihan Suplemen Kesehatan dalam Menghadapi Covid-19," J. Kesehat., vol. 13, no. 1, p. 65, 2022, doi: 10.26630/jk.v13i1.2814.

[9] F. Daryanes and D. Ririen, "Efektivitas Penggunaan Aplikasi Kahoot Sebagai Alat Evaluasi pada Mahasiswa," J. Nat. Sci. Integr., vol. 3, no. 2, p. 172, 2020, doi: 10.24014/jnsi.v3i2.9283.

[10] E. Ernawati, "Analisis Soal Ujian Matematika Berkategori Higher Order Thinking Skills (HOTS) Siswa SMP dengan Analisis Rasch Model," J. Didakt. Pendidik. Dasar, vol. 7, no. 1, pp. 277–294, 2023, doi: 10.26811/didaktika.v7i1.1027.

[11] R. Panji Yudha, "Pengembangan Instrumen Asesmen Otentik Tes Keterampilan Berpikir Kritis Pada Pembelajaran Matematika," M A T H L I N E J. Mat. dan Pendidik. Mat., vol. 4, no. 1, pp. 9–20, 2019, doi: 10.31943/mathline.v4i1.101.

[12] A. Faradhillah and C. Septiana, "MATHEMATICAL RESILIENCE: VALIDITY AND RELIABILITY WITH RASCH MODEL AND SPSS IN SENIOR HIGH SCHOOL," Sci. African, vol. 11. No 4, no. June, pp. 3545–3555, 2022, [Online]. Available: https://doi.org/10.24127/ajpm.v11i4.5204.

[13] N. S. R. Gumilang, W. Wahidin, and A. Tsurayya, "Pengembangan Instrumen Kemampuan Berpikir Kritis dan Kreatif Matematika Peserta Didik Kelas VII SMP," J. Pendidik. Mat. dan Sains, vol. 9, no. 2, pp. 89–98, 2021, doi: 10.21831/jpms.v9i2.27349.

[14] Y. D. Prastika, "Pengaruh Validitas, Reliabilitas dan Tingkat Kesukaran Terhadap Kualitas Butir Soal Ekonomi Menggunakan Software Anates di SMKN 3 Bangkalan," STKIP PGRI Bangkalan, pp. 1–11, 2021.

[15] C. Septiana and A. Faradillah, "Flipped Classroom Learning: Mathematical Creative Thinking Skills Based on Mathematical Resilience using Augmented Reality," J. Pendidik. Progresif, vol. 12, no. 3, pp. 1218–1231, 2022, doi: 10.23960/jpp.v12.i3.202217.

[16] M. Waruwu, "Pendekatan Penelitian Pendidikan: Metode Penelitian Kualitatif, Metode Penelitian Kuantitatif dan Metode Penelitian Kombinasi (Mixed Method)," J. Pendidik. Tambusai , vol. 7, no. 1, pp. 2896–2910, 2023.

[17] H. E. Nurmutia, "Pengaruh Gaya Kognitif Terhadap Kemampuan Pemecahan Masalah Matematis Siswa," Edumatika J. Ris. Pendidik. Mat., vol. 2, no. 2, p. 98, 2019, doi: 10.32939/ejrpm.v2i2.443.

[18] S. Maharani and M. Bernard, "Analisis Hubungan Resiliensi Matematik Terhadap Kemampuan Pemecahan Masalah Siswa Pada Materi Lingkaran," JPMI (Jurnal Pembelajaran Mat. Inov., vol. 1, no. 5, p. 819, 2018, doi: 10.22460/jpmi.v1i5.p819-826.

[19] Y. Handayani, R. Rahmawati, and ..., "Using Rasch Model to Analyze Reliability and Validity of Concept Mastery Test on Electricity and Magnetism Topic," JIPF (Jurnal Ilmu …, vol. 8, no. 2, pp. 226–239, 2023, [Online]. Available: https://journal.stkipsingkawang.ac.id/index.php/JIPF/article/view/3877.

[20] H. Herwin and S. C. Dahalan, "Person Fit Statistics to Identify Irrational Response Patterns for Multiple-choice Tests in Learning Evaluation," Pegem Egit. ve Ogr. Derg., vol. 12, no. 4, pp. 39–46, 2022, doi: 10.47750/pegegog.12.04.05.

[21] E. Hermansyah, C. A. Rizkyanti, and L. N. Hakim, "Rasch Calibration and Differential Item Functioning (DIF) Analysis of the School Well-Being Scale for Students," J. Kependidikan J. Has. Penelit. dan Kaji. Kepustakaan di Bid. Pendidikan, Pengajaran dan Pembelajaran, vol. 10, no. 1, p. 1, 2024, doi: 10.33394/jk.v10i1.10759.

[22] H. Mulyono, S. Kusuma Ningsih, and B. Sholeha Raufi, "Validating the Academic Writing Creativity and Self-Efficacy Scale: A Rasch Model Analysis," IRJE |Indonesian Res. J. Educ. |Vol, vol. 7, no. 1, pp. 121–135, 2023, [Online]. Available: https://doi.org/10.22437/irje