JIPF (JURNAL ILMU PENDIDIKAN FISIKA)
p-ISSN: 2477-5959 | e-ISSN: 2477-8451 *Vol. 9 No. 2, May 2024, Page 253-265*

Impressive Analysis of High School Students' Abilities in Physics Measurement Materials of the New Indonesian-Paradigm Curriculum Via the Rasch Model

Mohd Zaidi Bin Amiruddin [1*)], Achmad Samsudin [2], Bayram Coştu [3], Andi Suhandi [4], Ida Kaniawati [5], Nila Apriliyanti [6]
Universitas Pendidikan Indonesia, Indonesia[1,2,4,5], Yildiz Technical University, Turkey[3], SMA AL-Islam Krian Sidoarjo, Indonesia[6]
[*)]Corresponding E-mail: mohdzaidi@upi.edu

Received: September 27th, 2023. Revised: July 8th, 2024. Accepted: July 9th, 2024

Keywords :
Rasch Model; Merdeka Curriculum; Physics; Ability

ABSTRACT
The extent of students' abilities is an important thing to consider because it is part of the learning process. This research analyzes students' abilities in the subject matter measurement physics. Tools that are used is a ministep Rasch analysis. Descriptive quantitative was used in the study by providing an overview of the student's results and supported by data in the form of numbers from Rasch analysis. A total of 14 participant students as respondents in this research with the criteria of having received physics and measurement materialize Merdeka Curriculum in his learning. The results of this research state that ability learners are still classified as low, although it has obtained that material, while Cronbach's alpha index (0.62) is in the moderate category. Besides that, suitability questions for students still need to be classified as low. That way, it is important to balance the equations with students' abilities or vice versa. In the future, study continuation is required to increase students' abilities. According to the claim from the Merdeka Curriculum, it is learning more about the essentials.

INTRODUCTION

Measuring students' abilities in the field of education is very important in the learning process. This is useful for assessing the extent to which students have achieved learning goals and helping teachers and educational institutions design more effective curricula. In this context, measuring students' abilities is an important aspect of learning [1] [2] [3] [4]. According to Digna & Widyasari [5]; Halimat et al. [6]; Wahyuninsih & Lestari [7]; Zamista & Deswita [8], participant students' ability level will provide information to teachers to carry out appropriate education for students or what is known as "differentiated learning" in the Merdeka Curriculum. One lesson considered difficult is physics measurement material, which is an integral part of physics subjects at the high school level.

Through Rasch's analysis, this model can help teachers and schools learn and adopt policies that build each other together.

The Curriculum Merdeka is a new approach to education in Indonesia that aims to provide more freedom to educational institutions in designing curricula that suit the needs of students. In this context, measuring students' abilities becomes more important because a curriculum tailored to students' needs requires appropriate measurement tools to monitor and evaluate their achievements. Several ability measurements have been carried out on students, such as diagnostic tests [9] [10] [11] [12], formative tests [13] [14] [15] [16], and sub-summative tests or summative tests [17] [18] [19] [20]. However, the analysis used to determine students' abilities still uses conventional methods in mapping student profiles. One study that can help determine students' abilities is using the Rasch analysis model [21] [22] [23] [24].

Assessment plays an important role in efforts to advance education. In education, review is not only related to student learning outcomes but may be related to question, affective, and psychomotor assessments. Several studies that have been conducted previously stated that the level of difficulty is only based on operational verbs [25] [26] [27], based on expert validation [28] [29] [30], and student responses [31] [32] [33]. In this case, Rasch analysis has the potential to measure the level of difficulty of questions, reliability of questions, and students' abilities. Additionally, the current use of the curriculum in Indonesia requires mapping students' abilities to carry out differentiated learning. According to Engelhard [34] and Sumintono [35], the Rasch Model is a statistical method used to measure students' abilities. This model allows for more precise and objective measurements, considering the difficulty of each test item and the student's abilities. By applying the Rasch Model to the physics measurement material in the Merdeka Curriculum, we can get a better understanding of how far the participants are. Therefore, utilizing Rasch analysis in educational assessment with certain criteria is very important.

This background underlies the importance of research to measure the abilities of high school students in education, especially physics measurement material, using the Rasch Model in the context of the Merdeka Curriculum. The results of this research can provide valuable information to teachers, educational institutions, and educational policymakers about the effectiveness of the Rasch model in achieving learning objectives in the Merdeka Curriculum and provide insight for better curriculum development in the future. Apart from that, this research can also help understand whether there are differences in the abilities of students from various backgrounds or different schools so that it can help in designing more inclusive and equitable learning programs. The research questions (RQ) that will be answered in this research are as follows:

RQ 1. How is the distribution of student abilities and the level of difficulty of the questions?
RQ 2. How is the distribution of person-item fit statistics?
RQ 3. How is the summary statistic person-item reliability?

Theoretical Merdeka Curriculum
Merdeka Curriculum is a curriculum with intracurricular learning that has the specific aim of creating fun learning and education for students and teachers [36] [37] [38] [39]. According to the Ministry of Education, Culture, Research, and Technology, the Merdeka Curriculum introduced in February 2022 is one of the *Merdeka Belajar* (Learning) with a focus on essential material and developing the character of the Pancasila Student Profile. There are 6 dimensions to the Pancasila student profile such as: Belief, fear of God, and noble character, and kebhinekaan (diversity), cooperation, independence, critical thinking, creativity which can be seen in Figure 1 (a). Furthermore, for more details of the Pancasila Student Profile can be seen in Figure 1 (b).

The new paradigm in Indonesia is the implementation of the Merdeka Curriculum which was launched after the covid-19 outbreak. In the country of Indonesia, Pancasila is the basic ideology of the country so everything that is done is based on Pancasila [40] [41] [42]. The components in Figure 1 are

increasingly being researched and integrated in various ways to achieve the Profile of Pancasila learners who emphasize the essence of learning.

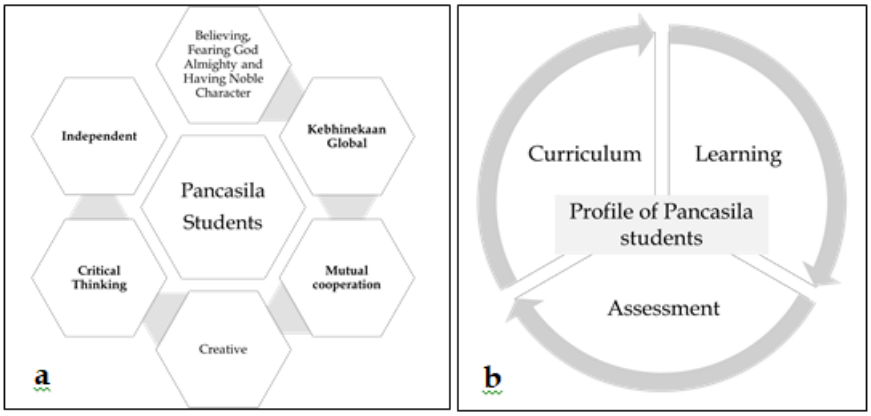


Fig 1. *The Framework merdeka curriculum*

The Merdeka Curriculum framework is increasingly being researched and used, especially in higher education institutions in Indonesia that produce prospective teachers. Implementing the Merdeka Curriculum, is not enough to only produce good teacher resources, it must be accompanied by adequate facilities to support learning [43] [44] [45]. In addition, the application of the Merdeka Curriculum continues to increase, as evidenced by the growth and decision-making of schools to implement it.

METHOD

Research design

This paper was used quantitative descriptive research methods to provide information about students' conditions objectively using responses obtained. This research used a descriptive quantitative approach because it was one type that can give a good interpretation, description, and appearance of results from the data obtained [46] [47] [48] [49]. The data obtained was then processed using Excel, ministep, and notepad. After that, it was described in accordance with what obtained from the results of data visualization.

Participant

The participants in this study were 14 students from a school in Sidoarjo, East Java. Each participant is a student who has just entered grade 11 and uses the Merdeka Curriculum. The participants also came from various backgrounds with the same ethnicity, namely the Javanese tribe. The following Figure 2 shows the location of the participants in this study.

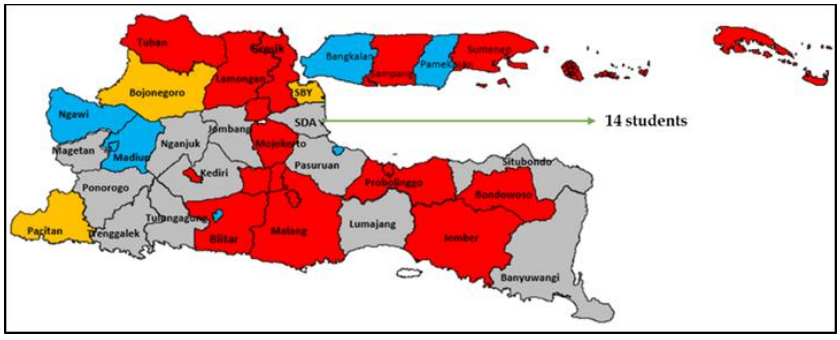


Fig 2*. The location of research participants*

Test instrument

The instrument in this research was ten multiple choice questions with physics measurement material. The ten questions are divided into several sections: derived basic quantities and dimensions, measuring area, measuring volume, rules for significant figures, and calculating speed. These questions are given to students online using the help of a "google form" with a predetermined time limit. Measurement material was chosen because it is the primary material in the Merdeka Curriculum for learning physics in Phase E. The following Figure 3 presents one of the sample problems used in this study.

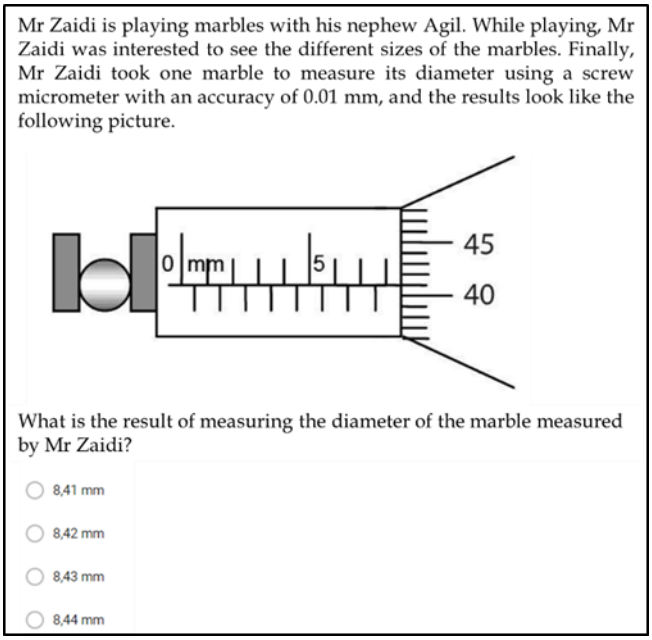


***Fig 3**. The example of one of the questions in the study*

Data Analysis

The data analysis process was done using Rasch analysis assisted by Rasch ministep software. The data analyzed are responses from students through a questionnaire given in multiple-choice form. The correct answer will be given a value of 1, and an incorrect answer will be 0. The data that has been obtained will be interpreted according to the following:

1. Wright Map (Person -Item Map). Distribution of students' abilities and distribution of difficulty levels of questions.
2. Level of suitability (Person-Item Fit). Data can be said to be fit when each Outfit mean square (MNSQ) value is in the range 0.5<MNSQ< 1.5 for and Mark Outfit Z-standard (ZSTD) is in the range -2.00< ZSTD< 2.00 and the point measure correlation (Pt Mean Corr) value is 0.4<Pt Measure<0.85 [50] [51]. The following is presented for the logit value categories in Table 1.

Table 1. Interpretation criteria

Item Logit Value (A)	Criteria	Person Logit Value (B)	Criteria
A ≥ +1.37	Very difficult	B ≥1.80	Height
0.0 ≥ A ≤ + 1.37	Difficult	B ≤ 1.80	Currently
0.0 ≤ A ≥ – 1.37	Currently	B ≤ -1,29	Low
A ≤ – 137	Easy		

3. Measuring value person and item reliability for its interpretation is presented as follows in Table 2.

Table 2. Interpretation of person and item reliability

Value Range	Interpretation
Value ≥ 0.95	Excellent
0.95 > value ≥ 0.91	Very Good
0.91 > value ≥ 0.81	Good
0.81 > value ≥ 0.68	Moderate
0.68 > value	Weak

4. Measure Reliability, namely the interaction between the person and the questions as a whole, with the following criteria in Table 3.

Table 3. Interpretation of Cronbach's Alpha values

Cronbach Alpha Range	Interpretation
α ≥ 0.8	Very Good
0.7> α ≥ 0.8	Good
0.6> α ≥ 0.7	Moderate
0.5> α ≥ 0.6	Bad
0.5> α ≥0.4	Very Bad

RESULTS

Distribution of Student Abilities and Level of Difficulty of Questions

Through Rasch analysis, we can find out the distribution of students' abilities and the difficulty level of the questions created. The distribution is presented below in Figure 4.

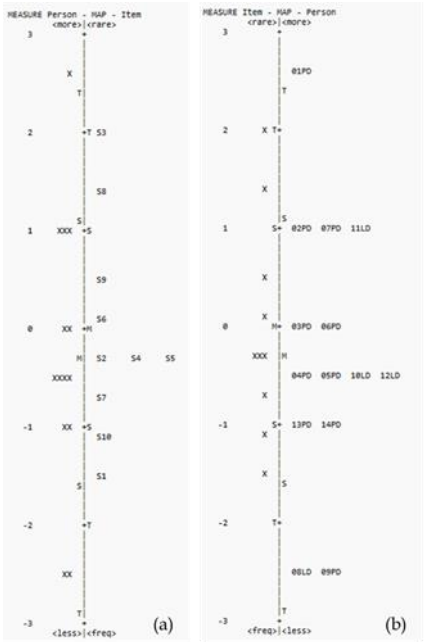


***Fig 4**. Map wright (Person -Item)*

Figure 4 presents two pictures related to item difficulty and student abilities. Figure 4 (a) discusses the questions' difficulty level based on students' ability to answer. From these data, it can be seen that the most difficult question was coded (S3) with a logit value of 2.04 (very difficult) and then followed by code questions (S8) and (S9) with logit values of 1.42 (very difficult) and 0.49 (difficult), respectively.

Apart from that, six questions have a logit value below zero, aka minus. This indicates that the question is in the easy category. According to Darmana et al. [52]; Fahmina et al. [53]; Lundström & Pesudovs [54], the logit value indicates the difficulty level of the question, where the higher the logit value, the more difficult the question will be. Figure 4 (b) discusses the level of students' abilities, which are measured based on their ability to answer the questions. In this case, students with the highest ability with the code (01PD) have a logit value of 2.61. Then, three students have the same ability as the code (02PD, 07PD, and 11LD) with a logit value of (1.00) in the medium category. Apart from that, students with a logit value ≤ -1.29 are in the low category.

Item Fit Order

The distribution of questions that fall into the fit or unfit categories needs to be considered in detail because it can influence the validity or invalidity of the questions asked. Below are presented the results of the Rasch analysis for the MNSQ, ZSTD, and Pt Measure Corr values in Figure 5.

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
2	7	14	-.28	.61	1.10	.48	1.92	1.79	A .34	.49	78.6	70.2	S2
9	5	14	.49	.64	1.43	1.40	1.18	.50	B .26	.48	50.0	75.1	S9
7	8	14	-.66	.62	1.23	.97	1.22	.59	C .34	.48	57.1	70.3	S7
6	6	14	.09	.62	1.12	.53	.95	.04	D .44	.49	57.1	72.2	S6
1	10	14	-1.48	.67	1.10	.40	.86	.04	E .42	.45	71.4	78.2	S1
8	3	14	1.42	.74	.89	-.12	.80	.04	e .51	.45	85.7	82.2	S8
10	9	14	-1.05	.64	.84	-.51	.66	-.49	d .59	.47	78.6	73.2	S10
4	7	14	-.28	.61	.80	-.80	.67	-.70	c .63	.49	78.6	70.2	S4
5	7	14	-.28	.61	.80	-.80	.66	-.74	b .63	.49	64.3	70.2	S5
3	2	14	2.04	.85	.63	-.62	.34	-.34	a .63	.41	92.9	87.4	S3
MEAN	6.4	14.0	.00	.66	1.00	.09	.93	.07			71.4	74.9	
P.SD	2.4	.0	1.02	.07	.23	.74	.41	.72			13.2	5.6	

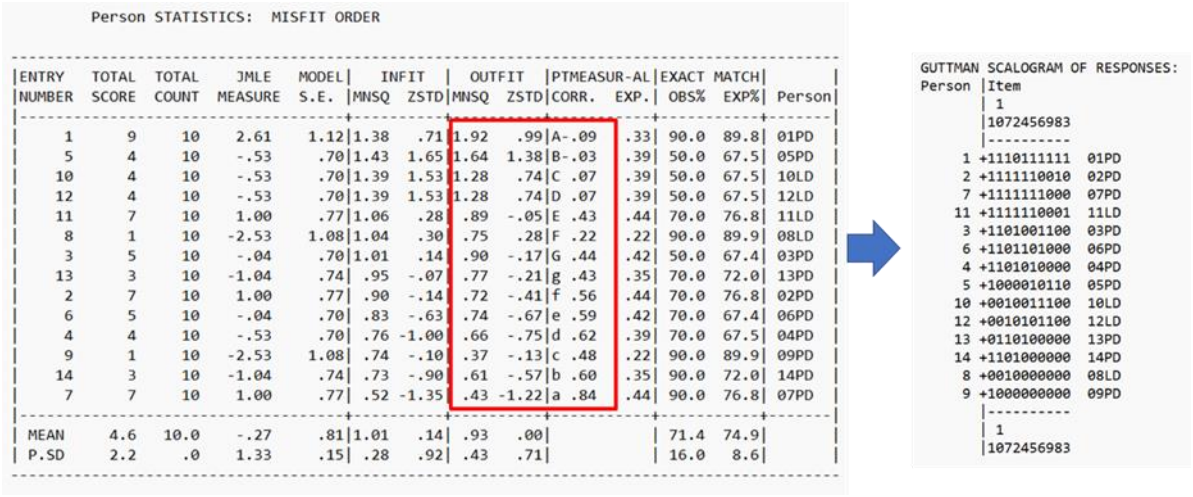
Fig 5. *Item measure for suitability of test items (validity)*

Figure 5 presents the suitability of questions that can use the MNSQ, ZSTQ, and Pt Measure Corr outfit parameters. Based on the information in Figure 3, it can be seen that all questions are included in the valid category, but some questions need improvement, such as questions (S2, S9, S7, and S3). These questions are said to need improvement because of the three parameters there is one that is not within the accepted range. For questions that are truly valid as a whole, namely questions (S6, S1, S8, SS10, SS4, and S5). According to Sumintono [35] a truly valid question is a question that has parameter sizes for all three that fall within the accepted categories or ranges that have been determined.

Person Fit Item

The suitability of students' abilities with the questions answered can also be one of the parameters for teachers in developing and providing difficulty for the questions. Apart from that, through the scalogram, we can also predict whether the answer was made by guessing, cheating, or luck. The following is a person measure fit in Figure 6.

Figure 6 presents information on person fit orders using MNSQ, ZSTD, and Pt Measure Corr criteria. In the person fit section, when one of the parameters is not appropriate, the student will be considered unfit so for all categories, they must be in the accepted category [55] [56] [57]. Person (01PD) has high abilities if seen from the logit value obtained. However, in this case, the question (S1) has a response pattern that is less suitable for students (01PD). Likewise, with other questions for other students. This can happen because the abilities and questions being worked on are not by [58] [59] [60]. To find out more, we can look at the Guttman scalogram.



Person STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE 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
1	9	10	2.61	1.12	1.38	.71	1.92	.99	A-.09	.33	90.0	89.8	01PD
5	4	10	-.53	.70	1.43	1.65	1.64	1.38	B-.03	.39	50.0	67.5	05PD
10	4	10	-.53	.70	1.39	1.53	1.28	.74	C .07	.39	50.0	67.5	10LD
12	4	10	-.53	.70	1.39	1.53	1.28	.74	D .07	.39	50.0	67.5	12LD
11	7	10	1.00	.77	1.06	.28	.89	-.05	E .43	.44	70.0	76.8	11LD
8	1	10	-2.53	1.08	1.04	.30	.75	.28	F .22	.22	90.0	89.9	08LD
3	5	10	-.04	.70	1.01	.14	.90	-.17	G .44	.42	50.0	67.4	03PD
13	3	10	-1.04	.74	.95	-.07	.77	-.21	g .43	.35	70.0	72.0	13PD
2	7	10	1.00	.77	.90	-.14	.72	-.41	f .56	.44	70.0	76.8	02PD
6	5	10	-.04	.70	.83	-.63	.74	-.67	e .59	.42	70.0	67.4	06PD
4	4	10	-.53	.70	.76	-1.00	.66	-.75	d .62	.39	70.0	67.5	04PD
9	1	10	-2.53	1.08	.74	-.10	.37	-.13	c .48	.22	90.0	89.9	09PD
14	3	10	-1.04	.74	.73	-.90	.61	-.57	b .60	.35	90.0	72.0	14PD
7	7	10	1.00	.77	.52	-1.35	.43	-1.22	a .84	.44	90.0	76.8	07PD
MEAN	4.6	10.0	-.27	.81	1.01	.14	.93	.00			71.4	74.9	
P.SD	2.2	.0	1.33	.15	.28	.92	.43	.71			16.0	8.6	

GUTTMAN SCALOGRAM OF RESPONSES:

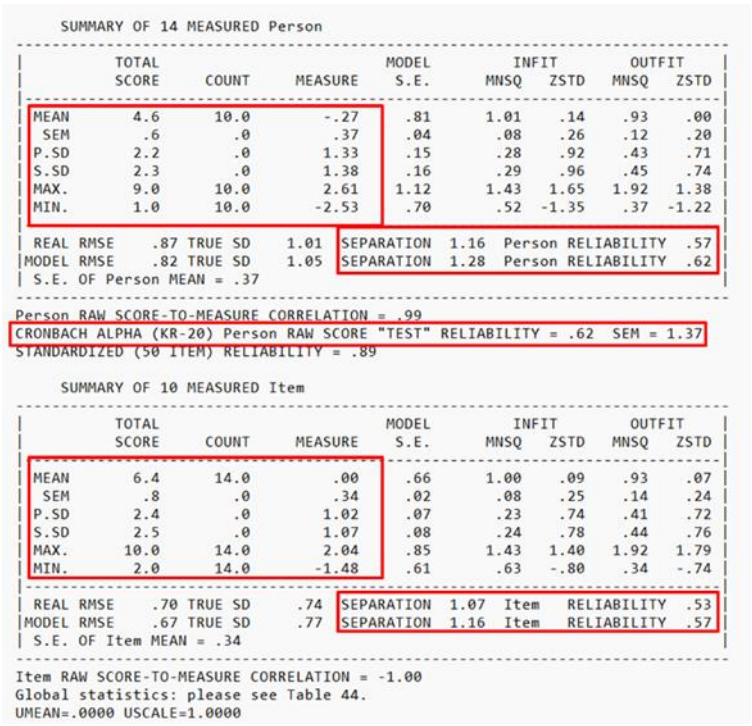
Person	Item 1	Person
	1072456983	
1	+1110111111	01PD
2	+1111110010	02PD
7	+1111111000	07PD
11	+1111110001	11LD
3	+1101001100	03PD
6	+1101101000	06PD
4	+1101010000	04PD
5	+1000010110	05PD
10	+0010011100	10LD
12	+0010101100	12LD
13	+0110100000	13PD
14	+1101000000	14PD
8	+0010000000	08LD
9	+1000000000	09PD
	1072456983	

Fig 6. *Person measure order with guttman scalogram*

Student 01PD has the highest ability with a logit value of 2.6 but cannot answer question number S2, which has a logit value of -0.28. This indicates that students may not be careful in answering these questions. According to Bouchée et al. [61]; Fatma [62]; Filighera et al. [63], several reasons why inappropriate response patterns occur are a lack of accuracy in answering questions and making guesses in answering. Based on the Guttman scalogram pattern, cheating behavior from each student has not been identified because they did not obtain the same type of answer.

Analysis Instrument Based on Summary Statistic

Summary statistics function and measurement information are displayed for overall instrument analysis. Summary statistics of the main menu are presented in Figure 7.



SUMMARY OF 14 MEASURED Person

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	4.6	10.0	-.27	.81	1.01	.14	.93	.00
SEM	.6	.0	.37	.04	.08	.26	.12	.20
P.SD	2.2	.0	1.33	.15	.28	.92	.43	.71
S.SD	2.3	.0	1.38	.16	.29	.96	.45	.74
MAX.	9.0	10.0	2.61	1.12	1.43	1.65	1.92	1.38
MIN.	1.0	10.0	-2.53	.70	.52	-1.35	.37	-1.22

REAL RMSE .87 TRUE SD 1.01 SEPARATION 1.16 Person RELIABILITY .57
MODEL RMSE .82 TRUE SD 1.05 SEPARATION 1.28 Person RELIABILITY .62
S.E. OF Person MEAN = .37

Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .62 SEM = 1.37
STANDARDIZED (50 ITEM) RELIABILITY = .89

SUMMARY OF 10 MEASURED Item

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	6.4	14.0	.00	.66	1.00	.09	.93	.07
SEM	.8	.0	.34	.02	.08	.25	.14	.24
P.SD	2.4	.0	1.02	.07	.23	.74	.41	.72
S.SD	2.5	.0	1.07	.08	.24	.78	.44	.76
MAX.	10.0	14.0	2.04	.85	1.43	1.40	1.92	1.79
MIN.	2.0	14.0	-1.48	.61	.63	-.80	.34	-.74

REAL RMSE .70 TRUE SD .74 SEPARATION 1.07 Item RELIABILITY .53
MODEL RMSE .67 TRUE SD .77 SEPARATION 1.16 Item RELIABILITY .57
S.E. OF Item MEAN = .34

Item RAW SCORE-TO-MEASURE CORRELATION = -1.00
Global statistics: please see Table 44.
UMEAN=.0000 USCALE=1.0000

Fig 7. *Summary statistic person-item reliability*

Based on Figure 7, it can be seen that the reliability aspect of students (0.57) and question items (0.53) are in the weak category, while Cronbach's alpha index (0.62) is in the moderate category. Based on

the information above, it can be seen that further study is needed regarding the reliability of students and the question items used because they are in the low category. That way, adjustments are required regarding the difficulty level of the questions and the level of students' abilities. On the other hand, this could mean that the evaluation is related to students' future development. The questions presented are part of the learning material that students have obtained. Abilities will continue to improve when they are often honed and trained consistently [64] [65].

DISCUSSIONS

Curriculum developments and changes indicate that policies are continuously being taken to improve the existing education system. In line with what was conveyed by Akkoyunlu [66]; Ratten [67]; Zhao & Watterston [68], curriculum changes were made to adapt educational needs to the needs of students and the demands of society. According to the Minister of Education, the curriculum used in Indonesia is the 2013 Curriculum and the Merdeka Curriculum. The implementation of the Merdeka curriculum still has certain obstacles. By implementing the Merdeka curriculum, there are no majors for senior high school at the beginning. Therefore, the learning carried out is still general.

One of the subjects given is physics. Physics learning begins with measurement material in the Merdeka Curriculum. According to Azita et al. [69]; Komalasari & Apriani [36]; Simamora & Pasaribu [70], the Merdeka Curriculum emphasizes the essence of learning. However, the results indicate that students' abilities are still relatively low when viewed from the person item map. This suggests that the learning that has taken place is not attached to the students even though the material being tested is part of the material that has been studied. With the Wright map, we can find out the student's abilities and the questions' difficulty level [21] [35] [71] [72]. It is important to continue to carry out assessments like this to provide a future picture of the level of ability, and the questions created must be appropriate so that students can develop.

According to Cayubit [73]; Gaol & Sitepu [74]; Mulang [75], students' abilities are influenced by several things, such as motivation, learning methods, school environment, and learning environment. Therefore, knowing and anticipating these factors is very important for learning to work well. No teaching is perfect, but it can be maximized with adequate preparation and facilities. Through summary item-person fit analysis, we can see the pattern of suitability of test items with students' abilities through logit values. The logit value between the question and the person can determine whether the student can answer a question; for example, when the logit value of the question is greater than the person's logit value, then logically, the student cannot answer the question [76] [77] [78] [79].

Many things must be considered regarding students' abilities in answering the physics measurement questions they have worked on. A total of 14 students answered these questions with different abilities. Cognitive learning alone is not enough for school because school is about learning and how comfort, character, and morals can be formed [80] [81] [82] [83]. Apart from that, students' interests in learning subjects also vary, so they require unique approaches and tests so that students can be directed according to their talents and interests.

CONCLUSION AND SUGGESTION

Based on the research results, it can be concluded that students' abilities, on average, are still in the medium category, as evidenced by the person logit value. Then, the question of reliability through Cronbach's alpha test is in the moderate category. In this way, it can be said that students' abilities are still lacking even though they have received physics measurement material, but there are still many errors in the answers. In the future, it is very important to implement differentiated learning so students' abilities can be directed and developed. So that the essence of education is evident in the student's abilities. Additionally, future research requires solutions and ways to improve student's

abilities so that they can provide breakthroughs in essential learning.

ACKNOWLEDGMENTS

The authors would like to thank the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia with DRTPM and "Program Pendidikan Magister Menuju Doktor untuk Sarjana Unggul (PMDSU) Batch VII" which has provided funding support and opportunities [Contract Number: 082/E5/PG.02.00.PL/2024].

REFERENCES

[1] Dede, C. (2010). Comparing frameworks for 21st century skills. *21st century skills: Rethinking how students learn*, *20*(2010), 51-76.
[2] Ango, M. L. (2002). Mastery of Science Process Skills and Their Effective Use in the Teaching of Science: An Educology of Science Education in the Nigerian Context. *Online Submission*, *16*(1), 11-30.
[3] Hartikainen, S., Rintala, H., Pylväs, L., & Nokelainen, P. (2019). The concept of active learning and the measurement of learning outcomes: A review of research in engineering higher education. *Education Sciences*, *9*(4), 276.
[4] Bernacki, M. L., Greene, J. A., & Crompton, H. (2020). Mobile technology, learning, and achievement: Advances in understanding and measuring the role of mobile technology in education. *Contemporary Educational Psychology*, *60*, 101827.
[5] Digna, D., & Widyasari, C. (2023). Teachers' Perceptions of Differentiated Learning in Merdeka Curriculum in Elementary Schools. *International Journal of Elementary Education*, *7*(2), 255-262.
[6] Halimah, N., Irdamurni, I., & Desyandri, D. (2023). Humanistic philosophy of learning differs from the curriculum in ES. *International Journal of Educational Dynamics*, *5*(2), 218-224.
[7] Wahyuninsih, S., & Lestari, Y. B. (2023). Implementation of Differentiated Instruction in Project Based English Language Learning: A Case Study at SMAN 1 Lambu. *International Journal of Multicultural and Multireligious Understanding*, *10*(9), 133-138.
[8] Zamista, A. A., & Deswita, P. (2023). Perceptions of Indonesian teachers on the implementation of "Merdeka" Curriculum. *Wiyata Dharma: Jurnal Penelitian dan Evaluasi Pendidikan*, *11*(1), 13-25.
[9] Adom, D., Mensah, J. A., & Dake, D. A. (2020). Test, measurement, and evaluation: Understanding and use of the concepts in education. *International Journal of Evaluation and Research in Education*, *9*(1), 109-119.
[10] Csapó, B., & Molnár, G. (2019). Online diagnostic assessment in support of personalized teaching and learning: The eDia system. *Frontiers in psychology*, *10*, 1522.
[11] Yuberti, Y., Suryani, Y., & Kurniawati, I. (2020). Four-tier diagnostic test with certainty of response index to identify misconception in physics. *Indonesian journal of science and mathematics education*, *3*(2), 245-253.
[12] Soeharto, S., Csapó, B., Sarimanah, E., Dewi, F. I., & Sabri, T. (2019). A review of students' common misconceptions in science and their diagnostic assessment tools. *Jurnal Pendidikan IPA Indonesia*, *8*(2), 247-266.
[13] Malespina, A., & Singh, C. (2022). Gender differences in test anxiety and self-efficacy: Why instructors should emphasize low-stakes formative assessments in physics courses. *European Journal of Physics*, *43*(3), 035701.
[14] Asriadi, M., & Hadi, S. (2021). Analysis of the Quality of the Formative Test Items for Physics Learning Using the Rasch Model in the 21st Century Learning. *JIPF (Jurnal Ilmu Pendidikan Fisika)*, *6*(2), 158-166.
[15] Kusairi, S. (2020). A web-based formative feedback system development by utilizing isomorphic multiple choice items to support physics teaching and learning. *Journal of Technology and*

Science Education, *10*(1), 117-126.
[16] Benben, V. Y., & Bug-os, M. A. A. C. (2022). Physics students' academic achievement and motivation in a gamified formative assessment. *American Journal of Educational Research*, *10*(6), 385-390.
[17] Albaiti, A., & Lepa, A. A. (2022). Solubility and Solubility Product Phenomena: Papua Senior High School Students' Mental Model. *Journal of Turkish Science Education*, *19*(2).
[18] Putra, A. (2020, March). Analysis of student's understanding about Newton's laws, in terms of perceptions to learning in senior high school. In *Journal of Physics: Conference Series* (Vol. 1481, No. 1, p. 012134). IOP Publishing.
[19] Pantiwati, Y., & Mahmudati, N. (2021). Increased ability to arrange two tier multiple choice questions. *Journal of Community Service and Empowerment*, *2*(1), 22-27.
[20] Kocimaheni, A. A., Soepardjo, D., Prasetyo, J., & Sasanti, N. S. (2023, March). Implementation of Kanji Learning Outcome Assessment During the Covid-19 Pandemic. In *International Joint Conference on Arts and Humanities 2022 (IJCAH 2022)* (pp. 984-991). Atlantis Press.
[21] Kania, V. I., Samsudin, A., Purwanto, A. H. A., Rasmitadila, R. R., Jermsittiparsert, K., & Nurtanto, M. (2020). Multitier of greenhouse effect (Moge) instrument development to identify middle school students' mental model in Thailand with rasch analysis,". *International Journal of Advanced Science and Technolo*, *29*(7), 3223-3237.
[22] Nasralla, M. M., Al-Shattarat, B., Almakhles, D. J., Abdelhadi, A., & Abowardah, E. S. (2021). Futuristic trends and innovations for examining the performance of course learning outcomes using the Rasch analytical model. *Electronics*, *10*(6), 727.
[23] Planinic, M., Boone, W. J., Susac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research: Why measurement matters. *Physical Review Physics Education Research*, *15*(2), 020111.
[24] Rahim, A., & Haryanto, H. (2021). Implementation of Item Response Theory (IRT) Rasch model in quality analysis of final exam tests in Mathematics. *Journal of Research and Educational Research Evaluation*, *10*(2), 57-65.
[25] Nurlailiyah, A., Deta, U. A., Ain, T. N., Haq, M. S., Lestari, N. A., & Yantidewi, M. (2019, February). Analysis of High School Physics National Examination questions based on Bloom Taxonomy and National Examination Question Standard in 2017/2018. In *Journal of Physics: Conference Series* (Vol. 1171, No. 1, p. 012041). IOP Publishing.
[26] Aprillia, A., Qadar, R., & Efwinda, S. (2023). Using Revised Bloom's Taxonomy to Evaluate the Cognitive Levels of Questions in Indonesian High School Physics Textbooks. *International Journal of STEM Education for Sustainability*, *3*(1), 195-211.
[27] Sole, F. B., & Anggraeni, D. M. (2020). Analysis of high order thinking skill (HOTS) in joint midterm examination at YAPNUSDA Elementary School. In *Journal of Physics: Conference Series* (Vol. 1440, No. 1, p. 012102). IOP Publishing.
[28] Istiyono, E., Dwandaru, W. S. B., Setiawan, R., & Megawati, I. (2020). Developing of Computerized Adaptive Testing to Measure Physics Higher Order Thinking Skills of Senior High School Students and Its Feasibility of Use. *European Journal of Educational Research*, *9*(1), 91-101.
[29] Wahyuni, S., & Husein, S. (2019, June). Physics learning devices based on guided inquiry with experiment to improve students' creativity. In *Journal of Physics: Conference Series* (Vol. 1233, No. 1, p. 012034). IOP Publishing.
[30] Bakri, F., Permana, H., Wulandari, S., & Muliyati, D. (2020). Student worksheet with ar videos: Physics learning media in laboratory for senior high school students. *JOTSE: Journal of Technology and Science Education*, *10*(2), 231-240.
[31] Cai, S., Liu, C., Wang, T., Liu, E., & Liang, J. C. (2021). Effects of learning physics using Augmented Reality on students' self-efficacy and conceptions of learning. *British Journal of Educational Technology*, *52*(1), 235-251.
[32] Hutchins, N. M., Biswas, G., Maróti, M., Lédeczi, Á., Grover, S., Wolf, R., ... & McElhaney, K. (2020). C2STEM: A system for synergistic learning of physics and computational thinking. *Journal of Science Education and Technology*, *29*, 83-100.
[33] Alam, A. (2020). Test of knowledge of elementary vectors concepts (TKEVC) among first-

semester bachelor of engineering and technology students. *Periódico Tchê Química*, *17*(35), 477-494.
[34] Engelhard Jr, G. (1992). The measurement of writing ability with a many-faceted Rasch model. *Applied measurement in Education*, *5*(3), 171-191.
[35] Sumintono, B. (2018, February). Rasch model measurements as tools in assesment for learning. In *1st International Conference on Education Innovation (ICEI 2017)* (pp. 38-42). Atlantis Press.
[36] Komalasari, M. D., & Apriani, A. N. (2023). Integration Of The Living Values Education Program (LVEP) In The Merdeka Curriculum. *Elementary School: Jurnal Pendidikan dan Pembelajaran ke-SD-an*, *10*(1), 61-69.
[37] Hadi, A., Marniati, M., Ngindana, R., Kurdi, M. S., Kurdi, M. S., & Fauziah, F. (2023). New paradigm of merdeka belajar curriculum in schools. *Al-Ishlah: Jurnal Pendidikan*, *15*(2), 1497-1510.
[38] Siregar, L. F., Sumanik, N. B., & Christianto, H. (2022). Analysis of teacher's ability in setting learning objectives, flow of learning objectives, and modules in the merdeka curriculum. In *SHS Web of Conferences* (Vol. 149, p. 01005). EDP Sciences.
[39] Pratama, S., Ramadhan, I., Hardiansyah, M. A., Wahyudi, A., Febrianty, G. S., & Alam, F. (2023, June). Merdeka Curriculum: Learning Systems and Challenges at Islamic Schools in Pontianak City. In *3rd Borneo International Conference on Islamic Higher Education (BICIHE) 2023* (Vol. 1, No. 1, pp. 90-96).
[40] Mukaromah, S. A., Gusmawan, A., & Munandar, J. (2022). The Lunge of Global Ideologies: The Challenges of Pancasila Ideology Education in the Middle of Global Existence in the Era of Globalization. *Jurnal Panjar: Pengabdian Bidang Pembelajaran*, *4*(1), 1-30.
[41] Suharto, T. (2018). Transnational Islamic education in Indonesia: an ideological perspective. *Contemporary Islam*, *12*(2), 101-122.
[42] Shofiana, G. F. (2014). Philosophy, Pancasila and modern technology. *Yuridika*, *29*(2).
[43] Pratikno, Y., Hermawan, E., & Arifin, A. L. (2022). Human resource 'Kurikulum Merdeka'from design to implementation in the school: What worked and what not in Indonesian education. *Jurnal Iqra': Kajian Ilmu Pendidikan*, *7*(1), 326-343.
[44] Sihombing, A. A., Anugrahsari, S., Parlina, N., & Kusumastuti, Y. S. (2021). "Merdeka Belajar" in an Online Learning during the COVID-19 Outbreak: Concept and Implementation. *Asian Journal of University Education*, *17*(4), 35-45.
[45] Tricahyati, S., & Zaim, M. (2023). English Teachers' Readiness in Implementation of 'Merdeka Belajar'Curriculum in Teaching English at Junior High School in Padang. *Journal of English Language Teaching*, *12*(1), 97-105.
[46] Eyisi, D. (2016). The usefulness of qualitative and quantitative approaches and methods in researching problem-solving ability in science education curriculum. *Journal of education and practice*, *7*(15), 91-100.
[47] Nehru, N., Ananda, R., Zb, A., & Navalyyan, D. (2022). The Analysis of Mathematical Critical Thinking Ability and Mathematical Creativity: Judging from the Process of Deriving the Fermi-Dirac Formula. *International Journal of Education and Teaching Zone*, *1*(2), 87-96.
[48] Nurkhin, A., & Pramusinto, H. (2020). Problem-Based Learning Strategy: Its Impact on Students' Critical and Creative Thinking Skills. *European Journal of Educational Research*, *9*(3), 1141-1150.
[49] Rodriguez, M. Y., & Storer, H. (2020). A computational social science perspective on qualitative data exploration: Using topic models for the descriptive analysis of social media data. *Journal of Technology in Human Services*, *38*(1), 54-86.
[50] Boone, W. J., Staver, J. R., & Yale, M. S. (2013). *Rasch analysis in the human sciences*. Springer Science & Business Media.
[51] Sumintono, B., & Widhiarso, W. (2014). *Aplikasi model Rasch untuk penelitian ilmu-ilmu sosial (edisi revisi)*. Trim Komunikata Publishing House.
[52] Darmana, A., Sutiani, A., Nasution, H. A., Ismanisa, I., & Nurhaswinda, N. (2021). Analysis of RASCH model for the validation of chemistry national exam instruments. *Jurnal Pendidikan Sains Indonesia (Indonesian Journal of Science Education)*, *9*(3), 329-345.
[53] Fahmina, S. S., Masykuri, M., Ramadhani, D. G., & Yamtinah, S. (2019, December). Content validity uses Rasch model on computerized testlet instrument to measure chemical literacy

capabilities. In *AIP Conference Proceedings* (Vol. 2194, No. 1). AIP Publishing.
[54] Lundström, M., & Pesudovs, K. (2009). Catquest-9SF patient outcomes questionnaire: nine-item short-form Rasch-scaled revision of the Catquest questionnaire. *Journal of Cataract & Refractive Surgery*, *35*(3), 504-513.
[55] Eddy, L. H., Preston, N., Mon-Williams, M., Bingham, D. D., Atkinson, J. M., Ellingham-Khan, M., ... & Hill, L. J. (2021). Developing and validating a school-based screening tool of Fundamental Movement Skills (FUNMOVES) using Rasch analysis. *PLoS One*, *16*(4), e0250002.
[56] An, M., & Yu, X. (2021). A Rasch analysis of emerging adults' health motivation questionnaire in higher education context. *PLoS One*, *16*(3), e0248389.
[57] Susilaningsih, E., Nuswowati, M., Aprilia, N., & Luthfiyah, A. (2021, June). Dissemination of test instruments as product of the development research to measure the problem-solving ability of class X students by online in the pandemic period. In *Journal of Physics: Conference Series* (Vol. 1918, No. 3, p. 032024). IOP Publishing.
[58] Husni, H. (2020). The effect of inquiry-based learning on religious subjects learning activities: An experimental study in high schools. *Jurnal Penelitian Pendidikan Islam*, *8*(1), 43-54.
[59] Chin, C., & Osborne, J. (2008). Students' questions: a potential resource for teaching and learning science. *Studies in science education*, *44*(1), 1-39.
[60] Ingber, D. E. (2020). Is it time for reviewer 3 to request human organ chip experiments instead of animal validation studies?. *Advanced Science*, *7*(22), 2002030.
[61] Bouchée, T., de Putter-Smits, L., Thurlings, M., & Pepin, B. (2022). Towards a better understanding of conceptual difficulties in introductory quantum physics courses. *Studies in Science Education*, *58*(2), 183-202.
[62] Ningsih, F. N. (2022). Analysis of Students' Concession Understanding Ability in Solving Physics Concepts. *International Journal of Education and Teaching Zone*, *1*(1), 26-33.
[63] Filighera, A., Ochs, S., Steuer, T., & Tregel, T. (2024). Cheating automatic short answer grading with the adversarial usage of adjectives and adverbs. *International Journal of Artificial Intelligence in Education*, *34*(2), 616-646.
[64] Lopes, P. N., Côté, S., & Salovey, P. (2013). An ability model of emotional intelligence: Implications for assessment and training. In *Linking emotional intelligence and performance at work* (pp. 53-80). Psychology Press.
[65] Alam, A. (2021, November). Possibilities and apprehensions in the landscape of artificial intelligence in education. In *2021 International Conference on Computational Intelligence and Computing Applications (ICCICA)* (pp. 1-8). IEEE.
[66] Akkoyunlu, B. (2002). Educational technology in Turkey: Past, present and future. *Educational Media International*, *39*(2), 165-174.
[67] Ratten, V. (2020). Coronavirus (Covid-19) and the entrepreneurship education community. *Journal of Enterprising Communities: People and Places in the Global Economy*, *14*(5), 753-764.
[68] Zhao, Y., & Watterston, J. (2021). The changes we need: Education post COVID-19. *Journal of educational change*, *22*(1), 3-12.
[69] Azita, A., Ersalianda, D., Andri, M., Mukhlis, M., & Tandon, M. (2023). Differences in the implementation of the 2013 curriculum and the independent learning curriculum at SMAN 3 Siak Hulu. *Lingeduca: Journal of Language and Education Studies*, *2*(2), 133-143.
[70] Simamora, R. M., & Pasaribu, D. (2023). Education should embrace all potential: students' reflective essays on the meaning of Merdeka Belajar. *Studies in Learning and Teaching*, *4*(1), 68-87.
[71] Aminudin, A. H., Kaniawati, İ., Suhendi, E., Samsudin, A., Coştu, B., & Adimayuda, R. (2019). Rasch analysis of Multitier Open-ended Light-Wave Instrument (MOLWI): Developing and assessing second-years sundanese-scholars alternative conceptions. *Journal for the Education of Gifted Young Scientists*, *7*(3), 557-579.
[72] Neumann, K., Viering, T., Boone, W. J., & Fischer, H. E. (2013). Towards a learning progression of energy. *Journal of research in science teaching*, *50*(2), 162-188.
[73] Cayubit, R. F. O. (2022). Why learning environment matters? An analysis on how the learning environment influences the academic motivation, learning strategies and engagement of college students. *Learning Environments Research*, *25*(2), 581-599.

[74] Gaol, R. L., & Sitepu, A. (2020). The Influence of Used Good-Based Learning Media on the Value of Chracter Education and Student's Motivation to Study. *Budapest International Research and Critics in Linguistics and Education (BirLE) Journal*, *3*(4), 1696-1703.
[75] Mulang, H. (2021). The effect of competences, work motivation, learning environment on human resource performance. *Golden Ratio of Human Resource Management*, *1*(2), 84-93.
[76] Nur, L., Nurani, L. A., Suryana, D., & Ahmad, A. (2020). Rasch model application on character development instrument for elementary school students. *International Journal of Learning, Teaching and Educational Research*, *19*(3), 437-459.
[77] Chan, S. W., Looi, C. K., Ho, W. K., Huang, W., Seow, P., & Wu, L. (2021). Learning number patterns through computational thinking activities: A Rasch model analysis. *Heliyon*, *7*(9).
[78] Chan, S. W., Looi, C. K., & Sumintono, B. (2021). Assessing computational thinking abilities among Singapore secondary students: A Rasch model measurement analysis. *Journal of Computers in Education*, *8*(2), 213-236.
[79] Edwards, A., & Alcock, L. (2010). Using Rasch analysis to identify uncharacteristic responses to undergraduate assessments. *Teaching Mathematics and its Applications: An International Journal of the IMA*, *29*(4), 165-175.
[80] Zeidler, D. L., Sadler, T. D., Simmons, M. L., & Howes, E. V. (2005). Beyond STS: A research-based framework for socioscientific issues education. *Science education*, *89*(3), 357-377.
[81] Zulela, M. S., Neolaka, A., Iasha, V., & Setiawan, B. (2022). How is the education character implemented? The case study in Indonesian elementary school. *Journal of Educational and Social Research*, *12*(1), 371.
[82] Lestari, F. P., Ahmadi, F., & Rochmad, R. (2021). The Implementation of Mathematics Comic through Contextual Teaching and Learning to Improve Critical Thinking Ability and Character. *European Journal of Educational Research*, *10*(1), 497-508.
[83] Prasetyo, T., Rachmadtullah, R., Samsudin, A., & Aliyyah, R. R. (2021). General Teachers' Experience of the Brain's Natural Learning Systems-Based Instructional Approach in Inclusive Classroom. *International Journal of Instruction*, *14*(3), 95-116.