JPPIPA 10(12) (2024)

Jurnal Penelitian Pendidikan IPA

Journal of Research in Science Education

http://jppipa.unram.ac.id/index.php/jppipa/index

Rasch Model: Analysis of Biology Question Item in the Indonesia Independent Curriculum

Kalika Azra Zainina[1], Mia Zubaidah Mufiqoh[1], Nani Aprilia[1*], Bekti Isnaeni[2]

[1] Biology Education, Faculty of Teacher Training and Education, Ahmad Dahlan University, Yogyakarta, Indonesia.
[2] Biology Education, Faculty of Mathematic and Science, Yogyakarta State University, Yogyakarta, Indonesia.

Received: January 30, 2024
Revised: August 14, 2024
Accepted: December 25, 2024
Published: December 31, 2024

Corresponding Author:
Nani Aprilia
naniapriliampd@gmail.com

DOI: 10.29303/jppipa.v10i12.7661

Abstract: The change from the 2013 curriculum to the independent curriculum has an impact on assessment, one of which is summative assessment. Summative assessment uses multiple choice questions, but the quality of the questions has not been analysed. This descriptive quantitative research aims to determine the quality of summative assessment questions in Biology class X odd semester which includes validity, reliability, difficulty level, and differentiating power. Research subjects were all grade X students. Object of research includes question grids, answer keys, questions, and student answer sheets. The sampling technique used was purposive sampling, and Kurmer Yogyakarta Pilot Project High School was selected. Data were collected through documentation. The instruments in this study were question grids, answer keys, questions, and student answer sheets. The data analysis technique used was quantitative descriptive analysis using Winsteps software version 3.73. The results of quantitative analysis of summative assessment questions in four schools showed valid results. The questions proved to be reliable, with difficulty levels in the easy and difficult categories. The questions had good and sufficient differentiating power. Distractor effectiveness was rated as poor, very good, good, and fair. Qualitative analysis of the 4 schools showed good results, but some improvements were still needed.

Keywords: Independent curriculum; Question items; Rasch model; Summative assessment

Introduction

The quality of education is needed to produce quality human resources so that all countries in the world are competing to improve the quality of education in their respective countries as well as in Indonesia. Through education, it is hoped that the realization of the objectives of national education, namely developing strong student character (Mahfud et al., 2024). One of the efforts to improve the quality of education in Indonesia is through the curriculum (Fatimah & Muamar, 2024; Khasana et al., 2024)). The curriculum in Indonesia has undergone several changes, until now the applicable curriculum is the independent curriculum (Ervia et al., 2024).

Curriculum development occurred with the change from the 2013 curriculum to the independent curriculum which was implemented in stages starting in the 2022/2023 school year (Fadilah et al., 2020; Hamdi et al., 2022). The independent curriculum is an emergency curriculum developed to overcome the learning crisis during the Covid-19 pandemic with the characteristics of learner-centered learning, focusing on instilling positive values in students and aiming to develop soft skills, understand concepts, and provide science exploration space for students (Fitria et al., 2024; Islamiati et al., 2024; Mardyah et al., 2024 ; Setyani et al., 2023). One of the main advantages of the independent curriculum is that learning is interactive, essential, flexible, and in-depth, in contrast to the previous curriculum approach which

How to Cite:
Zainina, K. A., Mufiqoh, M. Z., Aprilia, N., & Isnaeni, B. (2025). Rasch Model: Analysis of Biology Question Item in the Indonesia Independent Curriculum. *Jurnal Penelitian Pendidikan IPA*, *10*(12), 10990–10998. https://doi.org/10.29303/jppipa.v10i12.7661

only focused on essential material (Marjuni & Festiyed, 2024; Sukarno et al., 2024). The curriculum changes made by the government aim to improve the previous curriculum and provide flexibility to students and teachers to choose the most appropriate method according to the needs of their students (Nurhidayati et al., 2024).

This independent curriculum change has an impact on many things, one of which is learning (Mawati et al., 2023). Learning in the independent curriculum is by integrating a project model, one of which is with P5 activities or project activities to strengthen the profile of Pancasila students so that it can provide opportunities for students to be creative, active and independent (Kusyanti, 2022; Malikah et al., 2022; Saraswati et al., 2022). Learning with an independent curriculum is much more relevant and interactive for students to actively explore local, global, national, actual issues such as environmental, health, and other issues (Rahmadayanti & Hartoyo, 2022).

One of the challenges faced by teachers with the independent curriculum is that teachers find it difficult to evaluate learning even though evaluation is an integral part of learning (Alhayat & Riche, 2024; Poerwanti et al., 2024). Assessment is an important part of the learning process which consists of cognitive, affective, and psychomotor aspects (Swarni et al., 2024). Teachers need instruments to evaluate learning (Ariefiani & Laksono, 2024).

The assessment stages in the independent curriculum use an assessment approach, namely assessment as learning, assessment for learning and learning assessment. Assessment as learning is an assessment carried out between oneself and friends so that it involves students actively in assessment activities (Subehi & Sriyanto, 2021). Learning assessment is a continuous assessment process in the learning process such as assignments, presentations and projects (Rosana et al., 2020).

There are two types of assessment: formative assessment and summative assessment (Ismail et al., 2022). However, this research focuses on summative assessment. Summative assessment includes a variety of assessment formats with a multidisciplinary approach carried out in a structured manner with the aim of assessing learner learning outcomes (Buléon et al., 2022; Svensäter & Rohlin, 2023; Vittorini et al., 2021).

Based on the results of interviews with several biology subject teachers in several schools, information was obtained that the implementation of the independent curriculum in the odd semester of the 2022/2023 school year. The school has just implemented the independent curriculum for class X or stage E so that it becomes the object of this research. The material used in biology learning is more concise but more in-depth, in the odd semester there are 2 chapters, namely biodiversity and viruses, while in the even semester there is 1 chapter, namely environmental pollution which is a form of integration of biology, chemistry and physics subjects. This is due to the implementation of a project-based independent curriculum that is tailored to the themes in each school and learning includes intracurricular activities and projects that are linked to the Pancasila student profile.

The Rasch model is an analysis that can be used to measure the validity and reliability of research instruments, but it is also used to test the suitability of people and items simultaneously (Isnaeni et al., 2022). Analysis using the Rash model is easy to do and produces accurate measurements of student ability and question quality (Darmana et al., 2021; Suryandari et al., 2024). Rasch model analysis is an analysis that can convert data from ordinal format to ratio format. The data analyzed using the Rasch model are dimensionality, item validity, item ft, person reliability, item reliability, response choice order, and item invariance (Che Lah et al., 2021; Rencz et al., 2021; Silvia et al., 2021). The Rasch model also provides information on the measurement of the instrument's scale structure, so that the self-determination instrument is valid or invalid (Muslihin et al., 2022).

Winsteps is used to analyze Rasch model data by representing it on the same linear scale where each person is positioned on the same variable (Chan et al., 2021). Winstep software is a suitable measuring tool used in the Rasch model to measure item quality (Pohan et al., 2022). Winsteps is Windows-based software that can be used for Rasch model computing, especially in the fields of educational evaluation, attitude surveys, and scale analysis (Tarigan et al., 2022).

The novelty of this research is that there has been no research on the item analysis of the summative assessment of the independent curriculum at the Kurmer Pilot High School in Yogyakarta Special Region. The study aim was to analyze the quality of the summative assessment items of the independent curriculum at the Kurmer Pilot High School in the Special Region of Yogyakarta based on the aspects of validity, reliability, difficulty level, differentiating power, and effectiveness of triggers.

Method

Place, Time and Type of Research

This research was conducted in the Special Region of Yogyakarta at the high school pilot project of Kurmer, 4 schools were found, namely SMAN 7 Yogyakarta, MAN 2 Yogyakarta, SMAN 1 Pleret and MAN 3 Bantul.

The research was conducted from November 2022 to May 2023 of the 2022/2023 academic year. This research is included in quantitative descriptive research, Rasch model with the help of Winsteps software.

Research Subject and Object

The sampling technique used was purposive sampling, the selected school was the Kurmer pilot project in Yogyakarta. The research subjects were all grade X students who had implemented the independent curriculum in high schools selected as Kurmer pilot projects in Yogyakarta, totalling 931 students which can be seen at Table 1. The object of the research consists of the grids of summative assessment questions, answer keys to summative assessment questions, summative assessment questions and student answer sheets to summative assessment questions. The materials are biodiversity, viruses and environmental change.

Table 1. Totalling Student Class X

School	Class	Total Students
SMAN 7 Yogyakarta	X	284
MAN 2 Yogyakarta	X	247
SMAN 1 Pleret	X	179
MAN 3 Bantul	X	221

Data Collection Technique

The data collection technique used was documentation technique. The instruments used were documentation sheets in the form of question grids, answer keys to summative assessment questions, summative assessment questions and student answer sheets for summative assessment questions.

Data Analysis Technique

The data analysis technique in this study used quantitative descriptive analysis which included validity, reliability, level of difficulty, differentiating power and effectiveness of checkers. Data analysis was processed using the Rasch model assisted by Winsteps software version 3.73.

Result and Discussion

Quantitative Analysis
Validity

The validity of winstep software can be known through the item output: fit order which functions to analyse the suitability of items with certain criteria (Vaganian et al., 2022). Items that are declared fit (valid) or misfit (invalid) in the output adjust 3 conditions that must be met, namely the MNSQ Outfit value between 0.5 and 1.5; ZSTD outfit value between -2.0 and +2.0; and the measuring point correlation value between 0.4 and 0.85 (Azizah & Wahyuningsih, 2020). Validity is used to see if the items can measure what should be measured. A good test instrument must have a level of validity and reliability so that the test maker gets a correlation between the level of ability and difficulty of the questions in students (Tarigan et al., 2022).

Based on this description, the odd semester summative assessment questions at SMAN 7 Yogyakarta and SMAN 1 Pleret obtained valid results. Similarly, in MAN 2 Yogyakarta and MAN 3 Bantul, both type A and type B obtained valid results. This is because the highest percentage is in the valid category so it is said to be valid. The instrument is said to be valid according to the item fit criteria and there is no problem if the outfit values of MNSQ, ZSTD and Pt Measure Corr in the Person Fit Order output are in the range of 0.5 < MNSQ < 1.5; -2.0 < ZSTD < 2.0 and 0.4 < Pt Measure Correction < 0.80. The logit value is declared valid and can be used if at least two points of the above values are met (Ramadhan & Priatmoko, 2023). The results of the validity analysis can be seen in Figure 1.

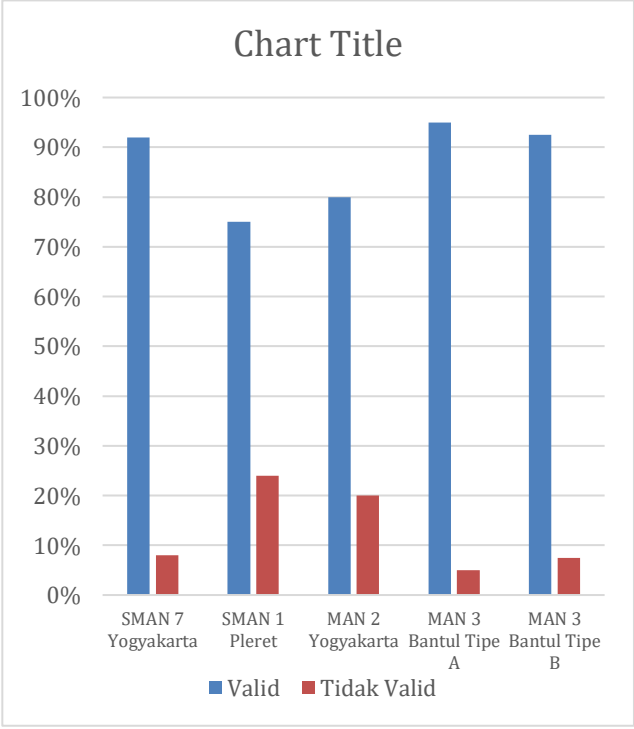


Figure 1. Odd Semester Summative Assessment Item Validity Results

Reliability

Reliability in Winstep software can be seen through the results of the output table in the statistical summary section which is then adjusted to the category, namely if the Item reliability value is more than 0.94 then it is included in the special category; if the value is between 0.91-0.94 it is included in the very good category; if the

value is between 0.80-0.90 it is included in the good category; if the value is between 0.67-0.80 it is included in the sufficient category and if the value is less than 0.67 it is included in the weak category (Handayani, 2022). Based on this description, the odd semester summative assessment questions at SMAN 7 Yogyakarta and SMAN 1 Pleret obtained excellent results. While MAN 2 Yogyakarta obtained excellent results and MAN 3 Bantul both type A and B obtained very good results. This shows that the school has compiled the items reliably or consistently. If the item reliability value is close to 1, then the reliability of the item is higher or better and the consistency or stability of the measurement results (Suryana et al., 2022). The results of the reliability analysis can be seen in Figure 2.

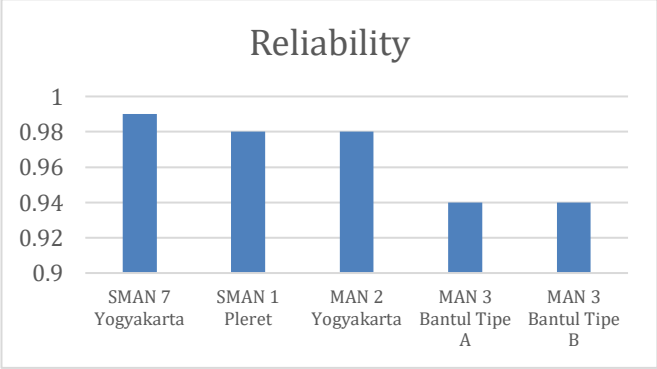


Figure 2. Results of Item Reliability of Odd Semester Summative Assessment Questions

Difficulty level

The level of difficulty through the Rasch model is seen from the logit size value and standard deviation (SD). The level of item difficulty can also be seen through the Wright map output where the left side of the Wright map is the level of item difficulty while the student's ability is on the right side of the Wright map (Chan et al., 2021; Chan et al., 2020). The difficulty level has 4 categories consisting of very easy, easy, difficult and very difficult categories. (Wulandari et al., 2022). The level of difficulty is a ratio or parameter that describes how difficult the test given to examinees is to give the correct answer (Suryani, 2017).

Based on this description, the highest percentage of odd semester summative assessment questions at SMAN 7 Yogyakarta is in the difficult category and SMAN 1 Pleret obtained the highest percentage result in the easy category. While MAN 2 Yogyakarta and MAN 3 Bantul type A obtained the highest percentage of difficult, besides MAN 3 Bantul type B obtained the highest percentage of easy. Test questions are said to be good if they have a moderate level of difficulty or not too easy and not too difficult (Arbiatin & Mulabbiyah, 2020). The results of the analysis of the level of difficulty of the questions can be seen in Table 2.

Table 2. Level of Difficulty of Odd Semester Summative Assessment Questions

School	Percentage %	Criteria
SMAN 7 Yogyakarta	46.10	Enough
MAN 2 Yogyakarta	45.00	Good
SMAN 1 Pleret	48.00	Good
MAN 3 Bantul Type A	57.50	Enough
MAN 3 Bantul Type B	45.00	Good

Differentiating power

Differentiating power in winstep software can be determined through item fit order, then see the correlation measuring point (Pt. Mean Corr) and adjusted to the criteria for differentiating power (Purwana et al., 2020). The differentiating power of items is the ability of item measuring instruments to distinguish students who have good mastery of competence from students who have not or lack mastery of competence based on certain criteria (Asrul et al., 2015).

Based on this description, the summative assessment questions of the odd semester of SMAN 7 Yogyakarta obtained the highest percentage results with the sufficient category and SMAN 1 Pleret obtained the highest percentage results with the good category. While MAN 2 Yogyakarta and MAN 3 Bantul type B obtained the highest percentage results with good category, besides MAN 3 Bantul type A obtained the highest percentage results with sufficient category. The results of the differentiating power analysis can be seen in Table 3.

Table 3. Differentiating Power of Odd Semester Summative Assessment Items

School	Percentage %	Criteria
SMAN 7 Yogyakarta	46.10	Enough
MAN 2 Yogyakarta	45.00	Good
SMAN 1 Pleret	48.00	Good
MAN 3 Bantul Type A	57.50	Enough
MAN 3 Bantul Type B	45.00	Enough

Effectiveness of distractors

The distractors are a pattern that can describe how students determine the answer choices to several answer choices (Sriyanto, 2019). Based on this description, the question of summative assessment of the odd semester of SMAN 7 Yogyakarta obtained the highest percentage result in the category of very good and SMAN 1 Pleret obtained the highest percentage result in the category of not good. While MAN 2 Yogyakarta and MAN 3 Bantul type A obtained the highest percentage of results with the good category, besides MAN 3 Bantul type B obtained the highest percentage of results with a fairly

good category. The results of the distractor effectiveness analysis can be seen in Table 4.

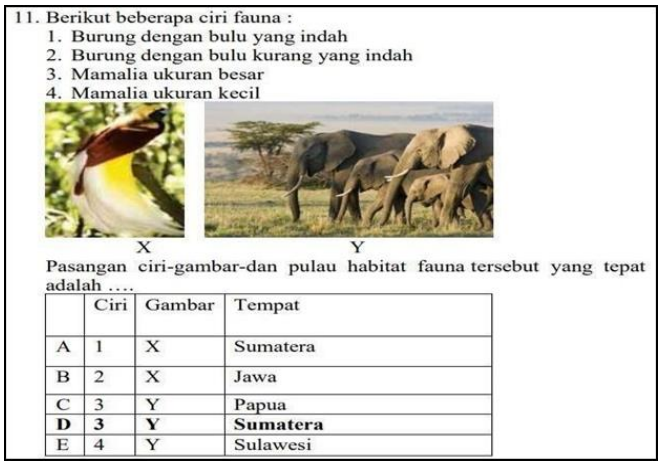
Table 4. Distractor Effectiveness of Odd Semester Summative Assessment Items

School	Percentage %	Criteria
SMAN 7 Yogyakarta	30.80	Very Good
MAN 2 Yogyakarta	27.50	Good
SMAN 1 Pleret	38	Deficient Good
MAN 3 Bantul Type A	35	Good
MAN 3 Bantul Type B	45	Good Enough

Qualitative Analysis

Qualitative analysis in this study used item review cards consisting of 3 aspects, namely material, construction and language. This analysis was carried out by 3 experts, namely researchers, biology teachers and peers. In addition, qualitative analysis was strengthened by Focus Group Discussion (FGD). FGD is a data collection technique in qualitative research methods to explore opinions, perceptions, and attitudes towards a concept or idea (Sugarda, 2020). FGD allows researchers and informants to discuss intensively for decision making, product or programme development, and knowing customer satisfaction (Fitriani & Azhar, 2019).

Based on the results of the qualitative analysis, it was found that at SMAN 7 Yogyakarta and SMAN 1 Pleret there were 27 questions that did not fit the construction aspect, 3 questions did not fit the language aspect and 6 questions did not fit the material aspect. In addition, at MAN 2 Yogyakarta and MAN 3 Bantul there were 16 questions that were not in accordance with the material aspect and 26 questions that were not in accordance with the construction aspect. The following are examples of items that are not in accordance with the construction aspect.


11. Berikut beberapa ciri fauna :
1. Burung dengan bulu yang indah
2. Burung dengan bulu kurang yang indah
3. Mamalia ukuran besar
4. Mamalia ukuran kecil

X Y

Pasangan ciri-gambar-dan pulau habitat fauna tersebut yang tepat adalah ....

	Ciri	Gambar	Tempat
A	1	X	Sumatera
B	2	X	Jawa
C	3	Y	Papua
D	**3**	**Y**	**Sumatera**
E	4	Y	Sulawesi

Figure 3. Example of a Summative Assessment Question Item that is Not in accordance with the Construction Aspect

Table 5. Qualitative Analysis Results

School	Material aspect	Construction aspect	Language aspect
SMAN 7 Yogyakarta	-	1, 2, 6	5, 7, 9
MAN 2 Yogyakarta	3, 7, 8, 24, 25, 37	-	1, 4, 6, 9, 10, 12, 16, 18, 20, 23, 25, 29, 32, 34, 35, 36, 39, 40, 43, 44, 46, 48, 49,50
SMAN 1 Pleret	18, 26	9, 12, 13, 19, 20, 28, 30, 31, 32, 33, 35, 36, 39	-
MAN 3 Bantul Type A	3, 12, 34	4, 35, 7, 11, 27, 18, 10, 28, 38	-
MAN 3 Bantul Type B	2, 4, 6, 7, 8, 9,10, 11, 12, 3, 18	1, 35	15, 36

Item number 11 does not comply with the construction aspect. This is due to the use of images in the form of elephants and statements on the question "large mammals", so that they can provide directions for the correct answer key. There are signs of preparing multiple choice questions, including not giving directions to the correct answer in the form of words, phrases or expressions (Afrian et al., 2017; Sumiati et al., 2018). Based on the results of qualitative analysis, there is a relationship with the results of quantitative analysis. The questions in the qualitative analysis results are not in accordance with the indicator "Questions do not provide answer key clues" in the construction aspect. Of course this can lead students to the correct answer without thinking, so that the choice of distractors or exemptions on the question items does not function. This is in accordance with the results of quantitative analysis on the aspect of the effectiveness of triggers which shows that the question is in the poor category. The function of distractors is to trick students into choosing the answers that are available (Putri & Rosliyah, 2020; Warju et al., 2020). The results of the qualitative analysis can be seen in Table 5.

Question Quality

A good test is a test that is able to distinguish the abilities of students, which students are learning and not learning, and does not show different results when used in different places. Good quality questions can also provide precise information on which students have not or have not understood the material that has been taught (Muluki et al., 2020). Questions that fall into the "Yes"

category can be directly included in the question bank because the quality of the questions is very good. Questions in the "No" category need to be revised so that they can be included in the question bank.

Questions with the category "No" means the quality of the item is poor, so the question should not be used or discarded. Based on this description, SMAN 7 obtained the results of items that entered the question bank as many as 8 items and SMAN 1 Pleret obtained the results of items that entered the question bank totalling 14 items while MAN 2 Yogyakarta obtained the results of items that entered the question bank as many as 20 items and MAN 3 Bantul type A obtained the results of items that entered the question bank totalling 25 items, besides MAN 3 Bantul type B obtained the results of items that entered the question bank totalling 26 items. The results of the question quality analysis can be seen in Table 6.

Table 6. Results of Quality Analysis of Summative Assessment Questions

School	Question Number	Criteria
SMAN 7 Yogyakarta	1, 2, 4, 5, 6, 7, 10 , 13	Yes
MAN 2 Yogyakarta	2, 3, 5, 10, 13, 14, 15, 16, 20, 21, 23, 25, 28, 29, 32, 34, 35, 37, 38, 40	Yes
SMAN 1 Pleret	2, 5, 11, 17, 19, 22, 26, 27, 33, 38, 41, 42, 45, 47	Yes
MAN 3 Bantul Type A	1, 2, 5, 6, 8, 9, 13, 14, 15, 16, 19, 20, 21, 22, 23, 24, 25, 26, 29, 30, 32, 33, 37, 39, 40	Yes
MAN 3 Bantul Type B	4, 5, 6, 7, 8, 10, 13, 14, 16, 17, 20, 21, 22, 23, 24, 25, 27, 28, 31, 32, 33, 34, 37, 38, 39, 40	Yes

Conclusion

The results of quantitative analysis of summative assessment questions in four schools showed valid results. The questions proved to be reliable, with difficulty levels in the easy and difficult categories. The questions had good and sufficient differentiating power. Distractor effectiveness was rated as poor, very good, good, and fair. Qualitative analysis of the 4 schools showed good results, but some improvements were still needed.

Acknowledgments
The author would like to thank the School as a research site and the Biology Education Study Program at Ahmad Dahlan University for facilitating this research.

Author Contributions
Conceptualization, K.A.Z, M.Z.M, and N.A.; writing original draft, methodology, investigation, software, result and disscussion, K.A.Z and M.Z.M; review, editing, format analysis and references, B.I. All authors have read and agreed to the published version of the manuscript.

Funding
This research received no external funding.

Conflicts of Interest
Regarding this study, the author declares that there is no conflict of interest.

References

Afrian, R., Islami, Z. R., & Mustika, F. (2017). Pembinaan Pembuatan Tes Buatan Guru (Soal) Mata Pelajaran Geografi SMA/MA Kota Langsa. *Jurnal Vokasi*, *1*(2), 108–116. https://doi.org/10.30811/vokasi.v1i2.687

Alhayat, A., & Riche, C. J. (2024). Designing Learning Assessment with STEAM-Based Projects in Implementing of the Kurikulum Merdeka. *Jurnal Penelitian Pendidikan IPA (JPPIPA)*, *10*(11), 9141–9157. https://doi.org/10.29303/jppipa.v10i11.7266

Arbiatin, E., & Mulabbiyah, M. (2020). Analisis Kelayakan Butir Soal Tes Penilaian Akhir Semester Mata Pelajaran Matematika Kelas VI SDN 19 Ampenan Tahun Pelajaran 2019/2020. *El Midad: Jurnal Jurusan PGMI*, *12*(2), 146–171. https://doi.org/10.20414/elmidad.v12i2.2627

Ariefiani, N. W., & Laksono, E. W. (2024). Development of An Integrated Assessment Instrument to Measure Students ' Critical Thinking and Chemical Literacy Skills for Rate of Reaction Topic. *Jurnal Penelitian Pendidikan IPA (JPPIPA)*, *10*(1), 7726–7734. https://doi.org/10.29303/jppipa.v10i10.9069

Asrul, A., Ananda, R., & Rosnita, R. (2015). *Evaluasi Pembelajaran*. Citapustaka Media.

Azizah, A., & Wahyuningsih, S. (2020). Penggunaan model rasch untuk analisis instrumen tes pada mata kuliah matematika aktuaria. *JUPITEK: Jurnal Pendidikan Matematika*, *3*(1), 45–50. https://doi.org/10.30598/jupitekvol3iss1pp45-50

Buléon, C., Mattatia, L., Minehart, R. D., Rudolph, J. W., Lois, F. J., Guillouet, E., Philippon, A. L., Brissaud, O., Lefevre-Scelles, A., Benhamou, D., Lecomte, F., Bellot, A., Crublé, I., Philippot, G., Vanderlinden, T., Batrancourt, S., Boithias-Guerot, C., Bréaud, J., de Vries, P., … Chabot, J. M. (2022). Simulation-based summative assessment in healthcare: an overview of key principles for practice. *Advances in Simulation*, *7*(1), 1–21. https://doi.org/10.1186/s41077-022-00238-9

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), 1–14. https://doi.org/10.1016/j.heliyon.2021.e07922

Chan, S. W., Looi, C. K., & Sumintono, B. (2020). Assessing computational thinking abilities among Singapore secondary students: a Rasch model measurement analysis. *Journal of Computers in Education*, *8*(2), 1–24. https://doi.org/10.1007/s40692-020-00177-2

Che Lah, N. H., Tasir, Z., & Jumaat, N. F. (2021). Applying alternative method to evaluate online problem-solving skill inventory (OPSI) using Rasch model analysis. *Educational Studies*, *49*(4), 1–23. https://doi.org/10.1080/03055698.2021.1874310

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*, *9*(3), 329–345. https://doi.org/10.24815/jpsi.v9i3.19618

Ervia, E., Harahap, R. D., & Chastanti, I. (2024). Analisis perkembangan kurikulum Biologi dari kurikulum 1984 sampai dengan kurikulum merdeka. *Didaktika: Jurnal Kependidikan*, *13*(1), 927–936. https://doi.org/10.58230/27454312.491

Fadilah, R., Parinduri, S. A., Syaimi, K. U., & Suharyanto, A. (2020). Islamic Guidance and Counseling to Overcome the Study Difficulty of Junior High School Students in SMP IT Nurul Azizi Medan (Case Study of Students experiencing Anxiety). *International Journal of Psychosocial Rehabilitation*, *24*(Special Issue 1), 1154–1160. https://doi.org/10.37200/IJPR/V24SP1/PR201262

Fatimah, F., & Muamar, M. R. (2024). Analysis of Students' Needs and Characteristics toward Science to Support Differentiated Learning of Kurikulum Merdeka. *Jurnal Penelitian Pendidikan IPA*, *10*(5), 2595–2602. https://doi.org/10.29303/jppipa.v10i5.6731

Fitria, D., Lufri, L., Asrizal, A., & Annisa, N. (2024). Digital Teaching Material of Integrated Science with Blended- PBL Model for Independent Curriculum. *Jurnal Penelitian Pendidikan IPA (JPPIPA)*, *10*(11), 8328–8338. https://doi.org/10.29303/jppipa.v10i11.9058

Fitriani, E., & Azhar, A. (2019). Layanan Informasi Berbasis Focus Group Discussion (FGD) dalam Meningkatkan Kepercayaan Diri Siswa. *Analitika: Jurnal Magister Psikologi UMA*, *11*(2), 82–87. https://doi.org/10.31289/analitika.v11i2.2552

Hamdi, S., Triatna, C., & Nurdin, N. (2022). Kurikulum Merdeka dalam Perspektif Pedagogik. *SAP (Susunan Artikel Pendidikan)*, *7*(1), 10–17. https://doi.org/10.30998/sap.v7i1.13015

Handayani, S. (2022). Analisis uas biologi kelas x dengan teori tes klasik dan item response theory (rasch model). *Bio-Pedagogi*, *11*(2), 76–90. https://doi.org/10.20961/bio-pedagogi.v11i2.63177

Islamiati, E. F., Subagia, I. W., & Suma, K. (2024). Development of Teaching Modules in the Implementation of the Independent Curriculum to Improve the Quality of the Learning Process and Student Learning Outcomes. *Jurnal Penelitian Pendidikan IPA (JPPIPA)*, *10*(11), 9097–9105. https://doi.org/10.29303/jppipa.v10i11.7521

Ismail, S. M., Rahul, D. R., Patra, I., & Rezvani, E. (2022). Formative vs. summative assessment: impacts on academic motivation, attitude toward learning, test anxiety, and self-regulation skill. *Language Testing in Asia*, *12*(1), 1–23. https://doi.org/10.1186/s40468-022-00191-4

Isnaeni, B., Aprilia, N., & Saifuddin, M. F. (2022). Rasch model: Quality of final semester assessment items for class x on biology subject. *Biosfer*, *15*(2), 355–367. https://doi.org/10.21009/biosferjpb.23326

Khasana, I., Viyanti, V., & Rosidin, U. (2024). Teachers ' Perception of Four-Tier Diagnostic Test Instrument Equipped with Self-Diagnosis Sheet to Identify Students ' Misconception on Linear Motion Material. *Jurnal Penelitian Pendidikan IPA (JPPIPA)*, *10*(9), 6910–6919. https://doi.org/10.29303/jppipa.v10i9.7720

Kusyanti, R. N. T. (2022). Analisis Standarisasi Laboratorium Fisika dalam Mendukung Implementasi Kurikulum Merdeka di SMA Negeri 1 Tempel. *Ideguru: Jurnal Karya Ilmiah Guru*, *8*(1), 40–47. https://doi.org/10.51169/ideguru.v8i1.404

Mahfud, H., Marmoah, S., Poerwanti, J. I. S., Sukarno, S., Supianto, S., & Istiyati, S. (2024). Implementation of Fostering an Attitude of Tolerance in Merdeka Curriculum on the Science Learning Process. *Jurnal Penelitian Pendidikan IPA (JPPIPA)*, *10*(11), 8454–8460. https://doi.org/10.29303/jppipa.v10i11.9316

Malikah, S., Winarti, W., Ayuningsih, F., Nugroho, M. R., Sumardi, S., & Murtiyasa, B. (2022). Manajemen Pembelajaran Matematika pada Kurikulum Merdeka. *Edukatif : Jurnal Ilmu Pendidikan*, *4*(4), 5912–5918. https://doi.org/10.31004/edukatif.v4i4.3549

Mardyah, U., Arif, K., Putri, A. N., Silvia, D., & Sari, Y. P. (2024). Misconception analisis of Junior High Scool Student in Lareh Sago Halaban Distrit on vibration, wave and sound materials using the four tier test instrument. *Jurnal Penelitian Pendidikan IPA (JPPIPA)*, *10*(10), 8197–8208. https://doi.org/10.29303/jppipa.v10i10.7629

Marjuni, M., & Festiyed, F. (2024). Development of Diagnostic, Formative and Summative Assessment Instruments in the PjBL Model to Strengthen the Profile of Pancasila Physics Students. *Jurnal Penelitian Pendidikan IPA (JPPIPA)*, *10*(10), 7520–7526. https://doi.org/10.29303/jppipa.v10i10.8758

Mawati, A. T., Hanafiah, H., & Arifudin, O. (2023). Dampak Pergantian Kurikulum Pendidikan Terhadap Peserta Didik Sekolah Dasar. *Jurnal Primary Edu (JPE)*, *1*(3), 69–82. https://doi.org/10.61116/jkip.v1i3.172

Muluki, A., Buntu, P., & Sukmawati, I. (2020). Analisis Kualitas Butir Tes Semester Ganjil Mata Pelajaran IPA Kelas IV Mi Radhiatul Adawiyah. *Jurnal Ilmiah Sekolah Dasar*, *4*(1), 86–96. https://doi.org/10.23887/jisd.v4i1.23335

Muslihin, H. Y., Suryana, D., Ahman, A., Suherman, U., & Dahlan, T. H. (2022). Analysis of the Reliability and Validity of the Self-Determination Questionnaire Using Rasch Model. *International Journal of Instruction*, *15*(2), 207–222. https://doi.org/10.29333/iji.2022.15212a

Nurhidayati, Firman, Erita, Y., & Daharnis. (2024). Development of Independent Curriculum IPAS Teaching Materials Using a Project-Based Learning Approach Assisted with the Kvisoft Flipbook Maker Application to Increase the Learning Motivation of Class IV Primary School Students. *Jurnal Penelitian Pendidikan IPA*, *10*(SpecialIssue), 279–285. https://doi.org/10.29303/jppipa.v10iSpecialIssue.7862

Poerwanti, J. I. S., Marmoah, S., Supianto, S., Sukarno, S., Mahfud, H., & Istiyati, S. (2024). Formative Assessment on Science Learning to Improve the Quality of Learning in Curriculum Merdeka. *Jurnal Penelitian Pendidikan IPA (JPPIPA)*, *10*(10), 7343–7353. https://doi.org/10.29303/jppipa.v10i10.9029

Pohan, L. A., Ledya, K., Setiawan, V. R., Darmana, A., & Suyanti, R. D. (2022). Penggunaan model rasch untuk mendeteksi kualitas soal kimia SMA pada ujian akhir semester (UAS). *EKSAKTA: Jurnal Penelitian Dan Pembelajaran MIPA*, *7*(2), 280–289. https://doi.org/10.31604/eksakta.v7i2.280-289

Purwana, U., Rusdiana, D., & Liliawati, W. (2020). Pengujian Kemampuan Menginterpretasikan Grafik Kinematika Calon Guru Fisika: the Polytomous Rasch Analysis. *ORBITA: Jurnal Kajian, Inovasi Dan Aplikasi Pendidikan Fisika*, *6*(2), 259–266. https://doi.org/10.31764/orbita.v6i2.3264

Putri, R. A., & Rosliyah, Y. (2020). Analisis Distraktor Butir Soal Bunpou Shokyuu Kouhan Semester 2 Tahun 2018/2019 Prodi Pendidikan Bahasa Jepang UNNES. *Kiryoku*, *4*(1), 18–25. https://doi.org/10.14710/kiryoku.v4i1.18-25

Rahmadayanti, D., & Hartoyo, A. (2022). Potret kurikulum merdeka, wujud merdeka belajar di sekolah dasar. *Jurnal Basicedu*, *6*(4), 7174–7187. https://doi.org/10.31004/basicedu.v6i4.3431

Ramadhan, Y. R., & Priatmoko, S. (2023). Pengembangan Tes Literasi Membaca dan Numerasi Materi Larutan Buffer Berbasis Kompetensi Minimum Untuk Identifikasi Kemampuan Metakognisi. *UNESA Journal of Chemical Education*, *12*(1), 51–58. https://doi.org/10.26740/ujced.v12n1.p51-58

Rencz, F., Mitev, A. Z., Szabó, Á., Beretzky, Z., Poór, A. K., Holló, P., Wikonkál, N., Sárdy, M., Kárpáti, S., Szegedi, A., Remenyik, É., & Brodszky, V. (2021). A Rasch model analysis of two interpretations of 'not relevant' responses on the Dermatology Life Quality Index (DLQI). *Quality of Life Research*, *30*(8), 2375–2386. https://doi.org/10.1007/s11136-021-02803-7

Rosana, D., Widodo, E., Setianingsih, W., & Setyawarno, D. (2020). Pelatihan Implementasi Assessment Of Learning, Assessment For Learning Dan Assessment As Learning Pada Pembelajaran IPA SMP di MGMP Kabupaten Magelang. *Jurnal Pengabdian Masyarakat MIPA Dan Pendidikan MIPA*, *4*(1), 71–78. https://doi.org/10.21831/jpmmp.v4i1.34080

Saraswati, D. A., Sandrian, D. N., Nazulfah, I., Abida, N. T., Azmina, N., Indriyani, R., Suryaningsih, S., Usman, U., & Lestari, I. D. (2022). Analisis Kegiatan P5 di SMA Negeri 4 Kota Tangerang sebagai Penerapan Pembelajaran Terdiferensiasi pada Kurikulum Merdeka. *Jurnal Pendidikan MIPA*, *12*(2), 185–191. https://doi.org/10.37630/jpm.v12i2.578

Setyani, A. I., Putri, D. K., Pramesti, R. A., Suryani, S., & Ningrum, W. F. (2023). Pembelajaran Biologi dalam kurikulum merdeka di Sekolah Urban. *DIAJAR: Jurnal Pendidikan Dan Pembelajaran*, *2*(2), 145–151. https://doi.org/10.54259/diajar.v2i2.1364

Silvia, P. J., Rodriguez, R. M., Beaty, R. E., Frith, E., Kaufman, J. C., Loprinzi, P., & Reiter-Palmon, R. (2021). Measuring everyday creativity: A Rasch model analysis of the Biographical Inventory of Creative Behaviors (BICB) scale. *Thinking Skills and Creativity*, *39*, 1–9. https://doi.org/10.1016/j.tsc.2021.100797

Sriyanto, A. (2019). Assesment Tes pada Hasil Belajar. *Al-Lubab: Jurnal Penelitian Pendidikan Dan Keagamaan Islam*, *5*(1), 117–142.

https://doi.org/https://doi.org/10.19120/al-lubab.v5i1.3732

Subehi, R., & Sriyanto, S. (2021). Implementasi Assessment Of, For, dan As Learning dalam Pembelajaran Daring PAI di SMPN 8 Purwokerto. *Alhamra Jurnal Studi Islam*, *2*(2), 111–122. https://doi.org/10.30595/ajsi.v2i2.10632

Sugarda, Y. B. (2020). *Panduan praktis pelaksanaan focus group disscussion sebagai metode riset kualitatif*. PT Gramedia Pustaka Utama.

Sukarno, S., Marmoah, S., Poerwanti, J. I. S., Istiyati, S., & Mahfud, H. (2024). Implementation of Independent Curriculum for Elementary Schools in Indonesian Schools Abroad. *Jurnal Penelitian Pendidikan IPA (JPPIPA)*, *10*(10), 7228–7235. https://doi.org/10.29303/jppipa.v10i10.8398

Sumiati, A., Widiastuti, U., & Suhud, U. (2018). Workshop Teknik Menganalisis Butir Soal dalam Meningkatkan Kompetensi Guru di SMK Cileungsi Bogor. *Jurnal Pemberdayaan Masyarakat Madani (JPMM)*, *2*(1), 136–153. https://doi.org/10.21009/jpmm.002.1.10

Suryana, D., Putri, M. A., Supriatna, M., & Yudha, E. S. (2022). Analisis rasch model: validitas dan reliabilitas instrumen korban bullying. *Hisbah: Jurnal Bimbingan Konseling Dan Dakwah Islam*, *19*(2), 199–214. https://doi.org/10.14421/hisbah.2022.192-12

Suryandari, K. C., Wijayanti, M. D., & Wahyudi, W. (2024). Using The Rasch Model to Measure Prospective Elementary Teachers' Perception toward Sciencepreneurship, Creative Thinking, Career Adaptability, and Personal Goals. *Jurnal Penelitian Pendidikan IPA (JPPIPA)*, *9*(7), 3975–3982. https://doi.org/10.29303/jppipa.v10i7.7381

Suryani, Y. E. (2017). Pemetaan kualitas empirik soal ujian akhir semester pada mata pelajaran Bahasa Indonesia SMA di Kabupaten Klaten. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *21*(2), 142–152. https://doi.org/10.21831/pep.v21i2.10725

Svensäter, G., & Rohlin, M. (2023). Assessment model blending formative and summative assessments using the SOLO taxonomy. *European Journal of Dental Education*, *27*(1), 149–157. https://doi.org/10.1111/eje.12787

Swarni, A., Herwin, H., & Sujati, S. (2024). Testing the construct validity and reliability of curiosity scale using confirmatory factor analysis (CFA). *Jurnal Penelitian Pendidikan IPA (JPPIPA)*, *10*(9), 6322–6330. https://doi.org/10.36941/JESR-2020-0080

Tarigan, E. F., Nilmarito, S., Islamiyah, K., Darmana, A., & Suyanti, R. D. (2022). Analisis Instrumen Tes Menggunakan Rasch Model dan Software SPSS 22.0. *JIPK: Jurnal Inovasi Pendidikan Kimia*, *16*(2), 92–96. https://doi.org/10.15294/jipk.v16i2.30530

Vaganian, L., Boecker, M., Bussmann, S., Kusch, M., Labouvie, H., Margraf, J., Gerlach, A. L., & Cwik, J. C. (2022). Psychometric evaluation of the Positive Mental Health (PMH) scale using item response theory. *BMC Psychiatry*, *22*(1), 1–11. https://doi.org/10.1186/s12888-022-04162-0

Vittorini, P., Menini, S., & Tonelli, S. (2021). An AI-Based System for Formative and Summative Assessment in Data Science Courses. *International Journal of Artificial Intelligence in Education*, *31*(2), 159–185. https://doi.org/10.1007/s40593-020-00230-2

Warju, W., Ariyanto, S. R., Soeryanto, S., & Trisna, R. A. (2020). Analisis Kualitas Butir Soal Tipe Hots Pada Kompetensi Sistem Rem Di Sekolah Menengah Kejuruan. *Jurnal Pendidikan Teknologi Dan Kejuruan*, *17*(1), 95–104. https://doi.org/10.23887/jptk-undiksha.v17i1.22914

Wulandari, T., Ramli, M., & Muzzazinah, M. (2022). Analisis Butir Soal Dynamic Assessment untuk Mengukur Pemahaman Konsep Klasifikasi Tumbuhan pada Mahasiswa. *Jurnal Pendidikan Sains Indonesia*, *10*(1), 191–201. https://doi.org/10.24815/jpsi.v10i1.22082