Al-Ishlah: Jurnal Pendidikan
Vol.16, 2 (June, 2024), pp. 1888-1896
ISSN: 2087-9490 EISSN: 2597-940X, DOI: 10.35445/alishlah.v16i2.4703

Evaluating Elasticity-Science Process Skills (E-SPS) in Construction Instrumen through Rasch Modelling

Diki Chen[1], Heru Kuswanto[2], Mundilarto[3], Ika Pratiwi[4], Muslimin Boma[5], Yovita Salsa Brilliant Hanindy[6]

[1] Universitas Negeri Yogyakarta, Indonesia; dikichen.2022@student.uny.ac.id
[2] Universitas Negeri Yogyakarta, Indonesia; herukus61@uny.ac.id
[3] Universitas Negeri Yogyakarta, Indonesia; mundilarto@uny.ac.id
[4] Universitas Negeri Yogyakarta, Indonesia; ikapratiwi.2022@student.uny.ac.id
[5] Universitas Negeri Yogyakarta, Indonesia; muslimin0046fip@student.uny.ac.id
[6] Universitas Negeri Yogyakarta, Indonesia; yovitasalsa.2022@student.uny.ac.id

ARTICLE INFO

Keywords:

Elasticity;
Science process skills
Rasch model;
Validation

Article history:

Received 2023-09-20
Revised 2024-02-06
Accepted 2024-05-28

ABSTRACT

The assessment of material elasticity science process skills in high school students requires validated and reliable instruments. This study aims to develop and validate a set of multiple-choice questions using the Rasch model to ensure their suitability for educational assessments. This quantitative research employed a model approach based on the Rasch model. The instrument consisted of multiple-choice questions designed to evaluate material elasticity science process skills. The validation process included two stages: content validation and construct validation. Content validation involved eight practicing teachers as validators, while construct validation was conducted with a sample of 32 high school students. Data analysis was performed using MNSQ infit, MNSQ outfit, delta, discrimination index, item differential power, and item difficulty level, with the assistance of Quest data analysis software. The results of the analysis from Quest revealed that only 10 questions met the Rasch model criteria. Among these, question number six was identified as the most difficult. The validation process demonstrated that the developed questions are appropriate for assessing the targeted skills based on the Rasch model. The findings indicate that the validated instrument is reliable and suitable for evaluating material elasticity science process skills in high school students. This study underscores the importance of rigorous validation processes in developing educational assessments. The developed multiple-choice questions, validated through the Rasch model, are feasible for use in assessing material elasticity science process skills. Future research should explore the application of this instrument in various educational contexts to further confirm its reliability and effectiveness across different student demographics.

Corresponding Author:
Diki Chen
Universitas Negeri Yogyakarta; dikichen.2022@student.uny.ac.id

1. INTRODUCTION

Physics is the study of natural phenomena explained through formulas that demonstrate these events (Astalini et al., 2021). In physics education, teachers typically introduce formulas, illustrate their applications in everyday life, and conduct laboratory experiments. However, a common issue in teaching physics is the ineffective classroom learning process (Yunzal & Casinillo, 2020; Maison et al., 2022). Students often remain passive, hindering the development of their affective skills. Additionally, laboratory sessions are infrequently utilized due to the extensive curriculum requirements. Conducting practical work necessitates support materials such as electronic learning kits (E-LKPD), rubrics, instrument questions, and various learning media. Despite their importance, educators often lack the time to adequately prepare these resources. This research aims to assist educators by developing instruments to assess science process skills related to elasticity.

Science process skills are essential techniques used by scientists to gather and interpret information about the natural world. These skills encompass a range of methods employed in the laboratory to uncover new knowledge and insights (Gunawan et al., 2019; Siregar et al., 2020). In educational settings, science process skills enable students to engage directly with scientific phenomena, fostering hands-on learning and critical thinking. Through activities that mirror the work of professional scientists, students can develop a deeper understanding of scientific concepts and methodologies. Science process skills include essential techniques that children use to acquire firsthand information through student learning activities. Effective science education incorporates these skills at every stage, ensuring that students actively participate in the learning process. The development of science process skills involves various student activities designed to enhance learning outcomes. Five key indicators of science process skills have been identified: observation, prediction, investigation, data interpretation, and communication.

Observation involves using the senses to gather information about objects or events. Prediction entails making educated guesses about future events based on current knowledge. Investigation includes designing and conducting experiments to test hypotheses. Data interpretation involves analyzing and making sense of collected data. Communication focuses on sharing findings and explanations with others. By integrating these indicators into science education, students can develop a comprehensive set of skills that not only enhance their understanding of science but also prepare them for future scientific endeavors.

In educational units, instrument questions are very important in the process of assessing and evaluating student performance. There are various types of questions such as; multiple choice, description, true-false, etc. The types of questions have different objectives, and the formulations are also different (Putranta & Supahar, 2019). The preparation of the instrument for science process skills is related to the indicators of the science process skills themselves. Indicators of observing, interpreting, predicting, and communicating must be in accordance with the questions and the choices of answers made. Questions that have good quality have a balance of difficulty levels (Rokhmat & Putrie, 2019; Bakri et al., 2019). Problems that are too easy do not stimulate students to increase their efforts to solve them. On the other hand, questions that are too difficult will cause students to become discouraged and not have the enthusiasm to try again because they are beyond their abilities. In its development, there are several models that can be used in test instruments, one of which is the Rasch model.

The Rasch model is an Item Response Theory (IRT) approach that considers item differences and item discriminants. IRT is a probability model that explains the relationship between a person's response to an item and the latent variable measured by the test (Hendriani & Septarini, 2016). The Rasch model is particularly useful for measuring the difficulty level of questions, analyzing the quality of items in an instrument, assessing student abilities, detecting biased questions, and identifying instances of cheating in creativity assessment instruments (Rahayu et al., 2021). Compared to other models, the Rasch model provides superior data measurement and accuracy, fulfilling five key principles: it provides linear data at equal intervals, predicts missing data, estimates data accurately, detects incorrect models, and produces replicable data (Planinic et al., 2019; Chan et al., 2021). In Rasch model analysis, item quality is assessed using the item information function to evaluate the overall quality of the questions and item fit to determine the quality of each individual question.

This research builds on previous studies, aiming to further the development of assessment tools for science process skills. Research by Hiğde and Aktamış (2022) developed a science process skills test instrument incorporating STEM, while Mufida et al. (2020) utilized a science process skills test integrated with Project-Based Learning (PjBL) and STEM. Similarly, Nunaki et al. (2020) integrated their science process skills test with an inquiry-based learning model. While these studies effectively focused on the learning models used to enhance science process skills, they did not fully explain the process of developing the test instruments themselves.

In light of these gaps, this research seeks to provide a detailed explanation of the development process of a multiple-choice question instrument specifically designed to assess elasticity science process skills. This study will address two primary questions: 1) What is the validity of the multiple-choice question instrument for assessing elasticity science process skills? and 2) How does the empirical test of the multiple-choice question instrument for elasticity science process skills perform based on the Rasch model? By answering these questions, this research aims to contribute a comprehensive and empirically validated tool for evaluating science process skills, thereby enhancing the accuracy and reliability of assessments in this field.

2. METHODS

This research aims to develop multiple-choice questions on the topic of elasticity, with a focus on validation and empirical testing based on science process skills indicators. This study adopts a quantitative research approach, utilizing the Rasch model for the analysis of the multiple-choice questions. The developed instrument comprises 12 multiple-choice questions designed to assess material elasticity science process skills across five key indicators.

The study subjects included eight practicing teachers who served as validators and a sample of 32 high school students who had previously studied elasticity. The indicators for the developed questions are detailed in Table 1. Validation and empirical testing were conducted to ensure the reliability and validity of the assessment tool, providing a robust measure of students' science process skills in the context of material elasticity.

Table 1. Distribution Questions Based on Indicators Science Process Skills

Indicators	Number of Questions
Observations	1, 2, 3, 4
Predictions	5, 6
Investigation	7
Data interpretation	8, 9, 10
Communications	11, 12

Two stages of validation were carried out, namely content validation and construct validation using the Quest software. To validate the questions, an assessor multiple choice questions is needed which is given to the validator. There are 3 aspects of the assessment with a total of 14 statements given to the validator. The range of scoring for the content validation stage can be shown in Table 2.

Table 2. Range Validation Categories Item Instrument Content

Range	Categories
45.51 – 56.00	Very good (no revision)
35.01 – 45.50	Good (needs a little revision)
24.51 – 35.00	Bad
14.00 – 24.50	Very bad

Then, for sample analysis using Quest software, the analysis variables taken are MNSQ infit, MNSQ outfit, disc, delta, item differential power, and item difficulty level.

3. FINDINGS AND DISCUSSION

The results obtained from the research must be supported by sufficient data to ensure their validity and reliability. The findings should address the research questions or hypotheses stated in the introduction. In this study, the results of data analysis include both content validation and construct validation, assisted by the Quest software.

Content validation involved eight practicing teachers as validators. These experts evaluated the relevance and appropriateness of the multiple-choice questions to ensure they effectively measured the targeted science process skills in elasticity. The results of this content validation process, as shown in Table 3, provide detailed insights into the feedback and recommendations from the validators. This thorough validation process is crucial for establishing the credibility of the instrument.

Construct validation, on the other hand, was conducted using empirical data from a sample of 32 high school students. This stage of validation analyzed the students' responses to the questions, utilizing the Rasch model to assess the quality and performance of each item. The Quest software facilitated this analysis by calculating key metrics such as item difficulty, discrimination indices, and fit statistics. These metrics help determine whether the questions accurately measure the intended constructs and whether they function consistently across different student populations.

The combination of content and construct validation ensures that the developed multiple-choice questions are both theoretically sound and empirically robust, providing a reliable tool for assessing science process skills related to elasticity.

Table 3. Content Validation Results

Validators	total score	Categories
WDA	56	very good
IP	48	very good
S	53	very good
MB	53	very good
SEA	51	very good
SAP	50	very good
AI	48	very good
RA	47	very good

Multiple choice question validation process will be declared complete if the score is in the very good category. Then, the next process is to carry out an empiricism test which aims to see whether the questions are appropriate or not, and the results can be seen in Figure 1.

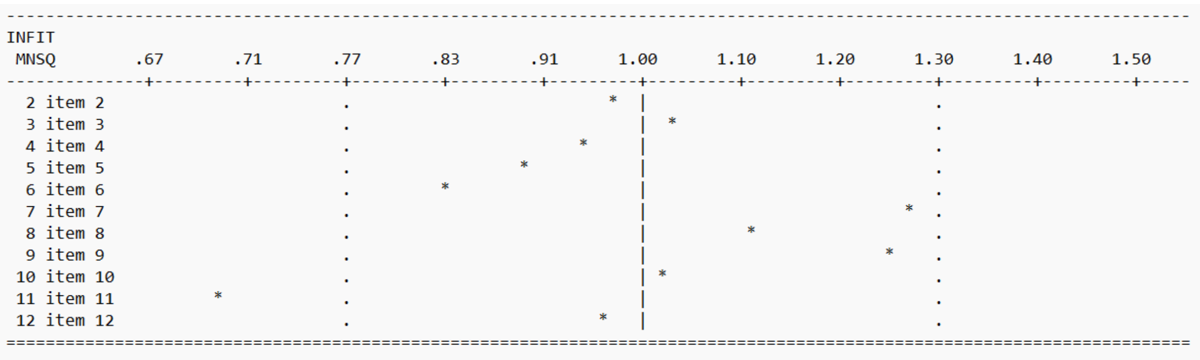

Figure 1. Output Rasch model mapping

A total twelve test items were tested on 32 samples, it was found that there were 10 statements that met the Rasch model criteria, questions number 1 and 11 did not meet the criteria because the MNSQ Infit score was not in. In full, the MNSQ Infit value for each item can be shown in Table 4.

Table 4. MNSQ Infit Values Per Item Questions

Item	Indicator			
	Infit MNSQ	Disc	Pt biserial (key answer)	Percentage key answer
1	0.00	0.00	0.00	100%
2	0.98	0.37	0.29	37.50%
3	1.03	0.29	-0.06	43.75%
4	0.94	0.41	-0.31	40.62%
5	0.89	0.40	0.40	84.37%
6	0.83	0.53	0.52	25%
7	1.27	-0.05	-0.05	37.50%
8	1.11	-0.11	-0.11	93.75%
9	1.25	-0.04	-0.04	59.37%
10	1.02	0.31	0.31	53.12%
11	0.70	0.73	0.72	56.25%
12	0.96	0.39	0.39	53.12%

From Table 4, the MNSQ Infit value shows the suitability of the analysis with the Rasch model. Items 1 and 11 do not meet the criteria of the Rasch model so it can be concluded that there are only 10 questions that are appropriate to use. The point-biserial value indicates the differentiability of the items. Then other analyzes raised by the Quest software can be seen in Figure 2.

```
------------------------------------------------------------------------------------------
    ITEM NAME       |SCORE MAXSCR|  DELTA |  INFT  OUTFT INFT  OUTFT
                    |            |    1   |  MNSQ  MNSQ   t     t
------------------------------------------------------------------------------------------
1   item 1          |     0     0 | Item has perfect score
                    |             |        |
2   item 2          |    12    32 |    .74 |    .98   .96   -.1   -.1
                    |             |    .39 |
3   item 3          |    14    32 |    .48 |   1.03  1.01    .3    .1
                    |             |    .38 |
4   item 4          |    13    32 |    .61 |    .94   .89   -.4   -.4
                    |             |    .38 |
5   item 5          |    27    32 |  -1.46 |    .89   .72   -.2   -.5
                    |             |    .50 |
6   item 6          |     8    32 |   1.34 |    .83   .74   -.8   -.7
                    |             |    .43 |
7   item 7          |    12    32 |    .74 |   1.27  1.38   1.8   1.4
                    |             |    .39 |
8   item 8          |    30    32 |  -2.43 |   1.11  1.66    .4    .9
                    |             |    .74 |
9   item 9          |    19    32 |   -.17 |   1.25  1.44   1.9   1.7
                    |             |    .38 |
10  item 10         |    17    32 |    .09 |   1.02  1.01    .2    .1
                    |             |    .38 |
11  item 11         |    18    32 |   -.04 |    .70   .67  -2.8  -1.5
                    |             |    .38 |
12  item 12         |    17    32 |    .09 |    .96   .92   -.3   -.3
                    |             |    .38 |
------------------------------------------------------------------------------------------
Mean                |             |    .00 |   1.00  1.04    .0    .1
SD                  |             |   1.08 |    .17   .32   1.3   1.0
==========================================================================================
```

Figure 2. Output data analysis using Quest

The first item was answered correctly by all students so there was no analysis on the pictures. Then the MNSQ outfit value is accepted if $0.5 < MNSQ < 1.5$, from the picture it can be seen that only the first and eleventh questions do not meet the criteria. Outfit t value ≤ 2.00 indicates that the sample fits the Rasch model, otherwise if Outfit t ≥ 2.00 then it does not fit the Rasch model.

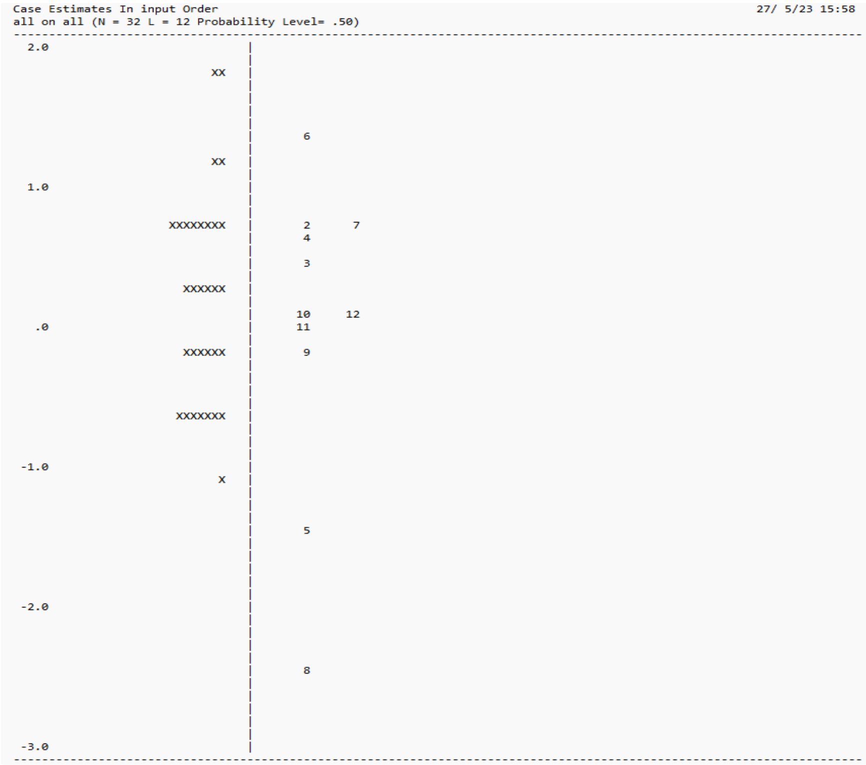

Figure 3. Output of difficulty level of the questions

Figure 3 shows the difficulty level of the questions, from easiest to most difficult. The minus sign indicates that the questions are relatively easy, whereas items 5 and 8 are relatively easy. Then, item 6 is a question that is very difficult for students to answer, followed by items 2 and 7, which have the same difficulty. Then a complete analysis of student abilities can be seen in Figure 4.

NAME	SCORE	MAXSCR	ESTIMATE	ERROR	INFIT MNSQ	OUTFT MNSQ	INFT t	OUTFT t
1 S01	6	11	.29	.66	.98	.85	-.06	-.10
2 S02	7	11	.73	.67	1.07	.94	.40	.13
3 S03	7	11	.73	.67	.95	.82	-.16	-.07
4 S04	4	11	-.61	.70	.83	.75	-.42	-.35
5 S05	3	11	-1.14	.76	1.24	1.28	.63	.61
6 S06	6	11	.29	.66	.76	.67	-1.22	-.51
7 S07	8	11	1.20	.71	.98	.77	.03	-.02
8 S08	4	11	-.61	.70	.68	.57	-.96	-.78
9 S09	5	11	-.14	.67	1.12	1.01	.54	.20
10 S10	4	11	-.61	.70	1.22	1.14	.71	.44
11 S11	7	11	.73	.67	1.06	.94	.33	.12
12 S12	7	11	.73	.67	1.23	3.51	1.08	2.60
13 S13	6	11	.29	.66	.88	.77	-.52	-.27
14 S14	5	11	-.14	.67	1.06	.97	.30	.10
15 S15	4	11	-.61	.70	.81	.79	-.47	-.25
16 S16	8	11	1.20	.71	.71	.56	-.99	-.36
17 S17	5	11	-.14	.67	1.04	1.08	.23	.32
18 S18	9	11	1.78	.81	1.15	1.13	.46	.48
19 S19	5	11	-.14	.67	1.37	1.35	1.39	.80
20 S20	6	11	.29	.66	1.48	2.81	2.12	2.43
21 S21	4	11	-.61	.70	.81	.79	-.47	-.25
22 S22	5	11	-.14	.67	.81	.72	-.70	-.44
23 S23	7	11	.73	.67	.83	.71	-.78	-.28
24 S24	7	11	.73	.67	1.21	1.08	.99	.34
25 S25	7	11	.73	.67	.89	.76	-.44	-.18
26 S26	4	11	-.61	.70	1.05	.90	.25	-.03
27 S27	6	11	.29	.66	1.07	1.26	.41	.63
28 S28	9	11	1.78	.81	1.15	1.13	.46	.48
29 S29	4	11	-.61	.70	.83	.75	-.42	-.35
30 S30	7	11	.73	.67	1.04	.89	.24	.05
31 S31	6	11	.29	.66	.73	.64	-1.40	-.57
32 S32	5	11	-.14	.67	.91	.81	-.29	-.24
Mean			.23		1.00	1.04	.04	.15
SD			.72		.20	.60	.77	.73

Figure 4. Output analysis of students' abilities

Figure 4 shows an analysis of students' abilities based on the number of samples studied, there are various kinds of analysis displayed, such as ability estimation, total number of questions answered correctly, and so on. Content validation and construct validation are needed to validate the questions made. Content validation is used to assess the content of both the material and the grammar used. Validators are needed who are experts in the field of education, such as lecturers or practitioners such as teachers/educators. Eight validators were used to assess the material elasticity science process skills multiple choice questions. The end result showed that the set of questions proposed was valid and ready to be tested for construct validity.

Construct validity is a picture that shows the extent to which the measuring multiple choice questions shows results that are in accordance with the theory. The process of testing construct validity is to link the measuring multiple choice questions with other measuring that have the same concept or with other measuring instruments that are theoretically related to it. There is a model that must be used when validating questions constructively. One of the data analysis techniques that can be used is the Rasch model. This model is a modern assessment theory that can group research instruments and subjects in a distribution map (Alavi et al., 2020). The Rasch model is a method that uses one parameter to reflect the test taker's abilities. Quest is one of the applications used to implement this model. Analysis using the Rasch model is relatively easy but produces accurate results. The Rasch model examines a test taker's probability of answering correctly on a test item with a dichotomous structure, by comparing the participant's ability with the item's level of difficulty (Nur et al., 2020; Darmana et al., 2021). For example, a participant has a 50% probability of answering an item correctly if the participant's ability is equivalent to the difficulty of the item. The suitability of the item analysis for the Rasch model this research can be seen in Figure 1, if the MNSQ infit value is in the range of 0.77 to 1.30, it can be said that the item is in accordance with the Rasch model. There are 10 questions that match the Rasch model, while the rest are not suitable because the scores do not meet the criteria. Then Table 2 shows some of the outputs of the analysis for each item, point-biserial shows the differential power of the correct answer choices selected, while the percentage of students choosing the correct option is lowest in question 6, namely 25% of the total 32 samples. Meanwhile, all students answered the first item correctly.

Other analyses can be seen in Figures 2 and 3, showing how many students answered the questions given correctly. The maximum score was obtained in the first question, where 100% of students answered correctly and the sixth question was the most difficult because only 8 out of 32 of the total sample answered. Figure 4 shows a complete analysis of student abilities, if the estimated error value is above 1.0 then the student is categorized as having high ability, the value $-1,0 < ER < 1,0$ then the student is categorized as having medium ability, and if the estimate error value is below 1.0 then the student is categorized as low ability. In general, there is only one student who has low ability, 4 students who have high ability, and the remaining 27 students are categorized as having medium ability. This shows that more or less the sample studied already has good abilities, but it should be noted that there was one student who only answered 4 questions correctly out of a total of 12 questions. Guidance from educators is needed so that these students can improve their abilities. Ways to do this include including these students in groups that adequately understand the material (van de Grift et al., 2019; Yang et al., 2020). The hope is that in the post-test after the practicum, the student will develop significantly in his abilities.

An innovative approach was taken in the development of this multiple-choice instrument, which combines interactive elements and real-world context to measure science process skills in a more holistic and in-depth manner. In contrast to conventional instruments, this instrument emphasizes the use of simulations, case-based questions, or virtual experiments that allow students to interact directly with the lesson material. Additionally, the integration of real-world context in the questions provides students with more interesting and relevant challenges, creating a dynamic and motivating learning environment (Chang et al., 2020; Ningtyas & Asri., 2021). Through this approach, this multiple-choice instrument aims to not only measure students' factual knowledge, but also encourage the development of analytical, problem-solving and critical thinking skills required in the scientific process. Thus, this instrument enriches students' learning experiences and increases the validity of measurements in measuring science process skills as a whole. In addition, this research can also be used as a literature reference for other researchers who will do the same thing. But on the other hand, there are some drawbacks of this research which can be used as suggestions for the development of further research. This research only explains the values/numbers obtained from the Quest analysis, which means that the research is of a quantitative type. For further development, it can be carried out based on an existing development model. This can cover the shortcomings of this research, which does not explain the stages of development scientifically.

4. CONCLUSION

Based on eight practitioners who were asked as validators, 12 questions were valid. In the construct validation, only 10 items were found that met the Rasch model criteria. Implications of this research are very significant in the educational context. The development of multiple-choice instruments that are more interactive and relevant to the real world can increase the effectiveness of evaluation of students' science process skills. With this approach, teachers can better measure students' analytical, problem-solving and creativity abilities in dealing with real situations. However, some weaknesses also need to be acknowledged. One is the possible difficulty in developing adequate questions with real-world contexts that are appropriate for all science topics. Additionally, significant time and effort is required to design questions that are interactive and varied. Recommendations for further research are to dig deeper into the integration of technology in instrument development, such as the use of applications or online platforms that enable the development and assessment of questions more efficiently. In addition, there needs to be collaboration between teachers, researchers and instrument developers to ensure that the questions developed are in accordance with the curriculum and learning objectives set. By addressing these weaknesses and following these recommendations, the development of multiple-choice instruments can be more effective in improving the measurement of students' science process skills.

Acknowledgments: We would like to thank for everyone who had support this research

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

Alavi, K., Isa, K., & Palpanadan, S. T. (2020). Application of Rasch Model on Resilience in Higher Education: An Examination of Validity and Reliability of Malaysian Academician Happiness Index (MAHI). *International Journal of Higher Education*, *9*(4), 261-271. https://doi.org/10.5430/ijhe.v9n4p261

Astalini, A., Darmaji, D., Kurniawan, DA, & Chen, D. (2021). Investigating student perceptions based on gender differences using e-module mathematics physics in multiple integral materials. *Journal of Indonesian Science Education* , *9* (4), 602-619. https://doi.org/10.24815/jpsi.v9i4.21297

Bakri, F., Marsal, O., & Muliyati, D. (2019). Textbooks equipped with augmented reality technology for physics topics in high school. *Journal of Research & Development of Physics Education* , *5* (2), 113-122. http://journal.unj.ac.id/unj/index.php/jpppf/article/view/12343

Chan, SW, Looi, CK, & Sumintono, B. (2021). Assessing computational thinking abilities among Singapore secondary students: a Rasch model measurement analysis. *Journal of Computers in Education* , *8* , 213-236. https://doi.org/10.1007/s40692-020-00177-2

Chang, S. C., Hsu, T. C., & Jong, M. S. Y. (2020). Integration of the peer assessment approach with a virtual reality design system for learning earth science. *Computers & Education*, *146*, 1-15. https://doi.org/10.1016/j.compedu.2019.103758

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. http://doi.org/10.24815/jpsi.v9i3.19618

Gunawan, G., Harjono, A., Hermansyah, H., & Herayanti, L. (2019). Guided Inquiry Model Through

Virtual Laboratory To Enhance Students' Science Process Skills On Heat Concept. *Journal of Educational Horizons* , *38* (2), 259-268. http://dx.doi.org/10.21831/cp.v38i2.23345
Hendriani, W., & Septarini, B. G. (2016). Pengembangan tes berpikir kritis dengan pendekatan item response theory. *Jurnal Penelitian dan Evaluasi Pendidikan*, *20*(1), 45-55. http://repository.unair.ac.id/id/eprint/95550
Hiğde, E., & Aktamış, H. (2022). The effects of STEM activities on students' STEM career interests, motivation, science process skills, science achievements and views. *Thinking Skills and Creativity* , *43* (1), 1-15. https://doi.org/10.1016/j.tsc.2022.101000
Maison, M., Kurniawan, DA, Wirayuda, RP, & Chen, D. (2022). Process Skills-Based E-Module: Impact On Analytical Thinking Skills. *Journal of Physics Education Research & Development* , *8* (1), 23-34. http://journal.unj.ac.id/unj/index.php/jpppf/article/view/21664
Mufida, SN, Sigit, DV, & Ristanto, RH (2020). Integrated project-based e-learning with science, technology, engineering, arts, and mathematics (PjBeL-STEAM): its effect on science process skills. *Biosphere: Journal of Biology Education* , *13* (2), 183-200. http://journal.unj.ac.id/unj/index.php/biosphere/article/view/15688
Ningtyas, L. E., & Asri, M. T. (2021). Profile of 10th grade high school students misconception on virus topic based on three-tier multiple choice diagnostic test. *Berkala Ilmiah Pendidikan Biologi (BioEdu)*, *10*(3), 674-682. https://doi.org/10.26740/bioedu.v10n3.p674-682
Nunaki, JH, Siagian, SIR, Nusantari, E., Kandowangko, NY, & Damopolii, I. (2020, March). Fostering students' process skills through inquiry-based science learning implementation. *Journal of Physics: Conference Series* , *1521* (4), 1-6. https://doi.org/10.1088/1742-6596/1521/4/042030
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. https://doi.org/10.26803/ijlter.19.3.24
Planinic, M., Boone, WJ, Susac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research: Why measurement matters. *Physical Review Physics Education Research* , *15* (2), 1-21. https://doi.org/10.1103/PhysRevPhysEducRes.15.020111
Putranta, H., & Supahar, S. (2019). Development of physics-tier tests (PysTT) to measure students' conceptual understanding and creative thinking skills: a qualitative synthesis. *Journal for the Education of Gifted Young Scientists* , *7* (3), 747-775. https://doi.org/10.17478/jegys.587203
Rahayu, W., Putra, MDK, Rahmawati, Y., Hayat, B., & Koul, RB (2021). Validating an Indonesian Version of the What Is Happening in This Class?(WIHIC) Questionnaire Using a Multidimensional Rasch Model. *International Journal of Instruction* , *14* (2), 919-934. https://doi.org/10.29333/iji.2021.14252a
Rokhmat, J., & Putrie, SD (2019). A strategy of scaffolding development to increase students' problem-solving abilities: The case of physics learning with a causalitic-thinking approach. *Journal of Turkish Science Education* , *16* (4), 569-579. https://doi.org/10.36681/tused.2020.8
Siregar, E., Rajagukguk, J., & Sinulingga, K. (2020). Improvement of science process skills using scientific inquiry models with algodoo media and quotient adversity in high school students. *Journal of Transformative Education and Educational Leadership* , *1* (2), 53-65. https://jteel.unimed.ac.id/index.php?journal=jteel&page=article&op=view&path%5B%5D=23
van de Grift, WJ, Houtveen, TA, van den Hurk, HT, & Terpstra, O. (2019). Measuring teaching skills in elementary education using the Rasch model. *School Effectiveness and School Improvement* , *30* (4), 455-486. https://doi.org/10.1080/09243453.2019.1577743
Yang, H., Su, J., & Bradley, KD (2020). Applying the Rasch model to evaluate the self-directed online learning scale (SDOLS) for graduate students. *International Review of Research in Open and Distributed Learning* , *21* (3), 99-120. https://doi.org/10.19173/irrodl.v21i3.4654
Yunzal, AN, & Casinillo, LF (2020). Effect of physics education technology (PhET) simulations: evidence from stem students' performance. *Journal of Education Research and Evaluation* , *4* (3), 221-226. https://doi.org/10.23887/jere.v4i3.27450