



Used of Rasch Model for Analysis of Students' Critical Thinking Skills Test Instrumens on Thermochemical Topics

Rizka Fahreza, Sandy Yudha, Rudi Purwanto, Retno Dwi Suyanti and Ayi Darmana

Master of Chemistry Education, Postgraduate Program, Universitas Negeri Medan, Medan-20221, Indonesia Correspondinge-mail: rizkafahreza5@gmail.com

Received: June 23th, 2022 Accepted: July 22th, 2022 Online Published: August 30th, 2022

Abstract: This study is descriptive study with quantitative approach. Sample determination by proportional sampling. The investigated variables are student abilities, which are reflected in end-of-semester exam results in chemistry courses. In this study, experimental method was used as a means of data collection. The test instrument uses a two-part multiple-choice model with a total of 30 items and five answer options. Using RASCH to evaluate models for testing students' critical thermal materials. This study was conducted to evaluate an experimental instrument on the critical thinking skills of SMA Negeri 1 Binjai Chemistry students. 25 students will participate in the workshop. The questions will be presented in multiple choice format, with a total of 30 questions on Thermochemistry. Rasch model is used to get the appropriate items. Analysis was performed using WinSteps software. The output of the Winsteps version 4.3.2 software found 12 incorrect items and 13 correct items, resulting in a final score of 13. Based on the analysis of write maps, students with 14 codes were known to have greater abilities or talents. 02, 04 and 18 have the lowest probability.

Keywords: fit, rasch model, winsteps, reliability, Critical Thinking Skills

Abstrak: Penelitian ini merupakan penelitian deskriptif dengan pendekatan kuantitatif. Penentuan sampel secara proporsional sampling. Variabel yang diteliti adalah kemampuan mahasiswa yang tercermin dari hasil ujian akhir semester pada mata kuliah kimia. Dalam penelitian ini, metode eksperimen digunakan sebagai alat pengumpulan data. Instrumen tes menggunakan model pilihan ganda dua bagian dengan jumlah 30 item dan lima pilihan jawaban. Menggunakan RASCH untuk mengevaluasi model untuk menguji bahan termal kritis siswa. Penelitian ini dilakukan untuk mengevaluasi instrumen eksperimen terhadap kemampuan berpikir kritis siswa Kimia SMA Negeri 1 Binjai. Sebanyak 25 siswa akan mengikuti workshop tersebut. Soal akan disajikan dalam format pilihan ganda, dengan total 30 soal tentang Termokimia. Model rasch digunakan untuk mendapatkan item yang sesuai. Analisis dilakukan dengan menggunakan software WinSteps. Output dari software Winsteps versi 4.3.2 ditemukan 12 item salah dan 13 item benar, sehingga menghasilkan skor akhir 13. Berdasarkan analisis menulis peta, siswa dengan 14 kode diketahui memiliki kemampuan atau bakat yang lebih besar. 02, 04 dan 18 memiliki probabilitas terendah.

Kata Kunci: fit, model rasch, winsteps, reliabilitas, Kerampilan Berfikir Kritis

• INTRODUCTION

In classical theory The difficulty of the questions depends on the abilities of the respondents. The list is not difficult (easy) for high-performing respondents. The list becomes more difficult for less experienced respondents. The interviewee's skill is shown

in elements that are not difficult (easy) Difficult questions indicate that respondents are less competent. The difficulty of the questions depends on the ability of the interviewees. even with the same grit The one with lower capacity is even heavier. And the one with greater capacity is lighter. Classical measurement theory (Classical test theory) cannot be used to compare the difficulty of an item with its ability to respond (Sudariono, 2012).

Qasem (2013) explains that one of the weaknesses of classical measurement theory is that this theory assumes that test scores that represent traits or abilities must be in a stable linear function, if an individual's score increases in the test the number of abilities must also increase. However, some high-ability individuals sometimes score low on tests, and the opposite may be the case for low-ability individuals.

A new measurement theory is needed to overcome the weaknesses of classical measurement theory. This study suggests the use of modern measurement theory or commonly called Item Response Theory (IRT) to analyze items so that students' abilities can be measured properly. Modern standard theory is called modern test theory. This theory focuses on grains, as opposed to classical theory. The IRT models each participant's response to each item in the exam. Items that include all types of items. Multiple choice questions that have wrong and correct answers (Novinda, et al., 2019).

Educational and psychological measurement theorists have developed a modern empirical theory, or item response theory, to alleviate the shortcomings of scientific test theory. There are four parameter-based IRT methods: 1-Parameter Logistic Model (1PL), a two-parameter logistic (2PL) model.), the logistic parameter model (3PL), the fourparameter logistic model (4PL) This study focused only on the 1PL model, or the main Rasch Rasch. One Direction Focus In the 1950s, the simple and elegant idea of observing the performance of elementary school students led George Rush to new discoveries. That is, the probability that the student answers the question correctly is similar to the student's performance. The complexity of the question.

In this study, measurements were carried out using a test instrument. This instrument is specifically designed to measure students' critical thinking skills on the topic of thermochemistry in chemistry subjects. The test instrument was designed by taking into account the construct component and the descriptive component. The construct component is that the items must provide an indication of the answers that will be obtained, from low to high stages. In this study, the questions will produce answers that show a certain qualitative range, namely two alternative answers, true or false. Hereinafter referred to as the dichotomous response pattern. The second component is a descriptive component, which is a component that explains several points on certain aspects. In this study, the aspect that is considered is how the students' ability to understand students' critical thinking skills on the topic of thermochemistry. Chemistry subjects are given to students majoring in Mathematics and Natural Sciences at SMA Negeri 1 Binjai. then analyzed using the Rasch model.

Many teachers still use the classical measurement model today, but this classical theory has several weaknesses. Another model that overcomes the weaknesses of the classical theory is the use of the Rasch measurement model (Chan, et al., 2013). MiniStep/WinStep software was used because the rash model uses a variable when analyzing candidate performance. Analysis using the Rash model is very simple but provides accurate analytical results. The assessment measures the probability of answering various questions correctly by comparing the student's ability to the difficulty of the question. Thus, if a student's ability is known to be equal to the difficulty level of the question, then 50% of the time the student will answer the question correctly (Dharmana, et al. 2021). According to Sumintono & Vidiaso (2015) and Linacre (2016), the Rasch model has several advantages. This means that the Rash model detects incorrect

responses, predicts results for missing data, and determines respondent merit. Such raw results can identify binomials and polynomials in the data and identify their structure, estimation, and evidence of fraud.

WinStep software is a Rasch model statistical tool used to analyze scores generated by test tools to determine MNSQ fit, ZSTD fit, factor size correlation, item reliability, and Kornbach alpha. If the square root mean of the approximate root mean is greater than the square root mean of 1 (one) fit. If the instrumental variance is greater than predicted by the Rasch model, if the inadequate value is less than 1, the instrumental variance is less than predicted by the Rasch model (Aziza and Saturday, 2020).

In this study, the researchers wanted to determine the thermal skill quality of students using the Rasch model method. This quality was evaluated according to various indicators, including the reliability of the Rasch model product. Therefore, it is considered that which one is suitable for the Rasch model and which one is not suitable for the Rasch model. In addition, WinStep software can calculate the Krobach alpha to determine the reliability of the product.

In the Rasch model, individuals are exposed to factors based on their abilities and level of difficulty. The probability of answering a task correctly is a function of the skill level related to the difficulty of the question. The main feature of the Rasch model is that it is unbiased and unambiguous. This model uses weight as the only relevant part of the test. Even in the Rasch model, the problem of selecting test items is fine-tuning the learning to the needs and goals of the learning.

• METHOD

This research is a descriptive research with a quantitative approach. Determination of the sample using proportional random sampling technique. The level to be studied is the student's ability according to the results of the final exam in chemistry. This study uses survey method to collect data. The test instrument uses a two-part multiple-choice format with a total of 30 items and five response options. This is used because the multiple choice form is more objective and reliable in seeing student responses, without being influenced by the subjectivity of the rater. Testing the validity of this study using the content validity test. The content validity of this test instrument is highly emphasized, so it is carried out in two ways, namely a qualitative study and then the results are processed using the Aiken (V) formula. The qualitative study used is based on the judgment of the expert with the criteria for review based on aspects of material, construction, and language. The qualitative study was conducted by five experts, including three secondary school chemistry teachers, one chemistry teacher and one measurement and evaluation expert (Lalio et al., 2022).

This study focuses on the analysis of test instruments using the Rasch model. This study involved 25 chemistry students at SMA Negeri 1 Binjai. Students are tested for their critical thinking skills using 30 challenging questions. Questions C4-C6 are evaluated. 1 is correct and 0 is correct when comparing multiple choice questions. Therefore, the resulting data is shared. Results The test results were analyzed using WinSteps software. The WinSteps program generates the appropriate parameter objects for the Rasch model. If the MNSQ matches, the correlation of the ZSTD values with the elements and a normal search indicates that a model-based intersection has been declared. This means that if the MNSQ value of the process is between 0.5 and 1.5; ZSTD time value from -2.0 to 2.0; and the correlation between units and quantities ranges from 0.4 to 0.85 (Sumintono & Vidhiarso, 2015, p. 12).

• RESULT AND DISCUSSION

By selecting Ministep in the Item Size parameter, you can monitor the item size complexity level or get the logit coefficient. Table 1 contains columns that provide information about the difficulty (item size) of each survey item. 13 items correspond to Rasch version with WinSteps software and 12 items do not correspond to Rasch.in. These results are presented throughout in

Table 1.

Table 1. Data on the Distribution of Misfit or Unfit Questions with the Rasch Model

ENTRY	TOTAL	TOTAL		MODEL	TA	IC T T		FETT	DIMEAS		EVACT	MATCHI	
NUMBER	SCORE	COUNT	MEASURE	S.E.	MNSO	ZSTD	MNSO	ZSTD	CORR.	EXP.	OBS%	EXP%	Person
5	6	25	-1.26	.49	1.19	.83	1.77	2.00	A15	.27	76.0	75.9	05
24	4	25	-1.80	. 56	1.12	.44	1.60	1.22	B03	.23	84.0	83.9	24
1	11	25	27	.42	1.34	2.43	1.34	2.03	C16	.41	36.0	64.1	01
23	11	25	27	.42	1.27	1.96	1.25	1.54	D05	.41	44.0	64.1	23
14	18	25	1.03	.47	1.12	.60	1.21	.78	E .10	. 29	72.0	73.1	14
15	10	25	45	.43	1.04	. 36	1.10	.63	F .22	.40	64.0	65.5	15
10	9	25	64	.44	.98	09	1.08	.43	G .29	. 30	80.0	67.2	10
18	2	25	-2.62	.75	1.03	.24	.98	.23	H .13	.47	92.0	92.0	18
13	5	25	-1.51	. 52	1.02	.17	.92	08	I .25	.45	80.0	79.9	13
16	7	25	-1.04	.46	1.00	.05	1.02	.15	J .27	.48	72.0	72.2	16
7	5	25	-1.51	. 52	1.01	.14	.98	.10	K .24	.25	80.0	79.9	07
19	10	25	45	.43	1.00	. 08	.97	11	L .31	. 30	64.0	65.5	19
4	2	25	-2.62	.75	.99	.18	.75	09	M .22	.47	92.0	92.0	04
9	14	25	.26	.42	.97	21	.99	. 00	1.35	.41	68.0	63.6	09
2	2	25	-2.62	.75	.97	.14	.89	.11	k .21	.47	92.0	92.0	02
21	10	25	45	.43	.97	18	.93	38	j.36	.40	64.0	65.5	21
20	3	25	-2.15	.63	.96	.05	.90	.03	1.26	.40	88.0	87.9	20
25	7	25	-1.04	.46	.96	11	.89	30	h .35	.48	72.0	72.2	25
8	3	25	-2.15	.63	.95	.04	.74	28	g .31	.40	88.0	87.9	08
22	13	25	.09	.42	.93	56	.91	58	f .42	.41	72.0	62.9	22
11	13	25	.09	.42	.91	72	.89	78	e .45	.31	64.0	62.9	11
17	10	25	45	.43	.84	-1.20	.80	-1.13	d .54	. 30	72.0	65.5	17
3	4	25	-1.80	. 56	.78	57	.54	-1.00	c .59	.23	84.0	83.9	03
12	3	25	-2.15	.63	.78	38	.47	92	Ь.57	. 20	88.0	87.9	12
6	14	25	.26	.42	.72	-2.33	.69	-2.26	a .71	.31	84.0	63.6	06
MEAN	7.8	25.0	-1.02	.51	.99	.1	.98	.1			74.9	74.8	
P.SD	4.5	.0	1.02	.11	.14	.9	.28	1.0	i		13.8	10.6	

Table 2 shows the distribution of errors or inconsistent cases according to the model. Element boundaries are described by the model only if one or both of the following conditions are met: The first condition is that the MNSQ score time must be between 0.5 and 1.5 (not according to code criteria 05, 24 and 12). Period ZSTD values between -2.0 and 2.0 (does not meet criteria for codes 01 and 06). and the items have correlation values between 0.4 and 0.85 in the total score (point measure correlation) (criteria code 05, 24, 14, 10, 07, 19, 11, 17, 17, 03, 12) and 06 is complete . Not). (Sumintono and Vidhyarso, 2014, pp. 71-72).

Table 1 shows the results of the test analysis using the WinSteps version 4.3.2 program. there are 12 items that are misfit, and 13 items of fit questions, so that the final instrument is 13 items.



Figure 1. Person Measure

Based on Figure 1 through the person measure, it can be concluded that the value of the student's ability level in working on the questions is shown from the Winsteps output, namely the Wright map. The data obtained is that students with code 14 have the highest ability or ability, while students 02, 04 and 18 have the lowest abilities.

• CONCLUSION

Similar to this Rasch model, the assessment instruments used in the accounting course on life insurance accounting were presented by extracting 12 items from the odd list and 13 items from the standard list, so that the final instrument was the result of Wright's study. 14 students with a package based on 13 items, S code is the best skill or ability. 02, 04 and 18 students have minimum skills.

• **REFERENCES**

- Arango, L. Dkk. 2022. Development of the German Social Attitude Barriers and Facilitators to Participation-Scales: an Analysis According to the Rasch Model. *Spinger link*. Vol 23(423).
- Asiah. N. Dkk. 2022. The Development and Validation of Indonesian Hope Scale (IHS): A Rasch Model Analysis. International Journal of Research in Counseling and Education. Vol 6(1).
- Astuti, B. Dkk. 2022. Scholl Connectedness Instrument's Testing with the Rasch Model for High Scholl Students During the COVID-19 Pandemic. *Cypriot Journal of Educational Sciences*. Vol 17(2): 410-421.
- Azizah dan Sapti, W. 2020. Penggunaan Model Rasch untuk Analisis Instrumen Tes Pada Mata Kuliah Matematika Aktuaria. *JUPITEK*. Vol 3(1): 45-50. 2020.
- Chan, S.W., Ismail, Z., & Sumintono, B. 2014. A rasch model analysis on secondary students' statistical reasoning ability in descriptive statistics. *Procedia-Social and Behavioral Sciences*,129:133-139Che, L.N.H., dkk. 2022. An Evaluation of the Online Social Learning Environment Instrument (OSLEI) Using Rasch Model Analysis. *SAGE Open*. Aril-June: 1-16.
- Darmana, A., dkk. 2021. Analysis of Rasch Model for the Validation of Chemistry National Exam Instruments. *Jurnal Pendidikan Sains Indonesia*. Vol 9(3): 329-345. 2021.
- Darmana, Ayi. Dkk. 2013. Pandangan Siswa Terhadap Internalisasi Nilai Tauhid melalui Materi Termokimia. *Prosiding Seminar FMIPA Universitas Lampung*.
- Harahap, A. Dan Darmana, A. 2020. Pembelajaran PBL Menggunakan Bahan Ajar Terintegrasi Nilai Spiritual untuk Meningkatkan Hasil Belajar dan Sikap Spiritual. Vol 2(2): 64-70.
- Jahro, I. S., dkk. 2021. Improving Students Science Processs and Critical Thinking Skills Using Semi-Research Patterns Practicum. *Jurnal Tadris Kimiya*. Vol 6(1):82-91.
- Laliyo, L. Dkk. 2022. Measuring Changes in Hydrolysis Concept Of Students Taught by Inquiry Model: Stacking and Racking Analysis Techniques in Rasch Model. *Heliyon*. Vol 8(3).
- Linacre, J.M. 2016. A User's Guide to WINSTEPS MINISTEP Rasch-Model Computer Programs, IL: Winsteps.com, Chicago.
- Liu, F. Dkk. 2022. Assessing the Psychometruc Properties of the Chinese Return-to-work Self Efficacy Questionnaire Using Rasch Model Analysis. *Spinger*. Vol 20(27).
- Masito, F. Dkk. Scientific Argumentation Skills Through the Rasch Model on Analysis of Survey Data on the Importance of Aviation Vocational Education in Indonesia. *Journal of Innovation in Educational and Cultural Research*. Vol 3(3).
- Mulyanti, S. Dkk. 2022. Validity Analysis Using the Rasch Model In The Dvelopment of Alkane Concept Test Instruments. *Jurnal Penelituan Pendidikan IPA*. Vol 8(3).

- Novinda, M., dkk. 2019. Pengembangan Tes Pilihan Ganda Menggunakan Model RASCH Materi Gerak Lurus Kelas X Pontianak. Artikel FKIP Untan Pontianak.
- Qasem, M. A. N, (2013). Journal of Research and Method in Education. A Comparative Study of Classical Theory (Ct) and Item Response Theory (Irt) In Relation To Various Approaches of Evaluating the Validity and Reliability of Research Tools. Volume 3(5): 77-81.

Sudaryono. (2012). Dasar-Dasar Evaluasi Pembelajaran. Yogyakarta: Graha Ilmu.

- Sumintono, B. & Widhiarso, W. 2015. Aplikasi Pemodelan Rasch pada Assessmen Pendidikan, Trim Komunikata, Cimahi.
- Syafitri, Anis dan Darmana, Ayi. 2018. Development of Chemistry Module Integrated With Islamic Values in Thermochemistry and Reaction Rate for Senior High School Student. *Jurnal Pendidikan Kimia*. Vol 10(3): 418-423.