Jurnal Pembelajaran Sains
VOLUME 9 NOMOR 1, JUNE 2025

http://journal2.um.ac.id/index.php/
ISSN: 2527-9157

UTILIZING RASCH MODEL AS NOVEL ANALYSIS OF STUDENTS' CRITICAL THINKING IN PHYSICS

Hilman Qudratuddarsi[1,*)], Jumriani[1], Meili Yanti[1], Nurhikma Ramadhana[1], Hefi Rusnita Dewi[2]
[1]Science Education Department, Faculty of Teacher Training and Education, Universitas Sulawesi Barat, Majene, Sulawesi Barat, Indonesia
[2]Mathematics Education, STKIP PGRI Bangkalan, Bangkalan, Jawa Timur, 69116, Indonesia

Email*: hilman.qudratuddarsi@unsulbar.ac.id

Abstract

There are numerous reported studies to find the best way of teaching physics, but their quantitative analysis merely depends on some inferential statistics comprising ANCOVA, t-tests, etc. Those analyses do not generate detailed information of student's abilities; therefore, this study analyzes the effects of double-loop problem-solving teaching methods towards student's critical thinking using Rasch analysis. In this study, the samples were divided into an experimental group and a control group with 28 students and 25 students, respectively. After the intervention, their critical thinking was measured using an instrument that has been validated and pilot tested using Rasch analysis by considering their reliability, item fit statistics, and unidimensionality. After data analysis, based on person measure, it was found that there is a statistically significant post-test logit score between the experiment and control class with pre-test as covariate, p value <0.05. This result is also supported by Wright-map which revealed the relationship between item and respondent at single picture. In the map, it can clearly be seen that experiment group has a higher logit compared to the control group. Based on differential item functioning (DIF) analysis in Rasch model, analysis on each item revealed that no significant difference in some items, moreover some items are more preferable to control groups. This deeper analysis can give new insight on the way of interpreting research data to select preferable teaching methods.

Keywords: *Rasch Model, Critical Thinking, Data Analysis, Statistical analysis*

Received: 29 December 2025 | Revised: 8 March 2025 | Published: 28 June 2025

INTRODUCTION

The studies to find the best teaching strategies will always be striking educational research due to the need to update teaching method to meet current condition (Ha et al., 2020). For instance, the world some years ago faced covid-19 pandemic and forces the learning from face-to-face meeting to online learning (Atmojo et al., 2020). Therefore, it is pivotal to put extra efforts how to teach science subjects effectively. Science teaching also adapted the world development by integrating science with mathematics, technology and engineering, commonly referered to STEM education (Apkarian et al., 2021). There are a plethora of studies (e.g (Bogusevschi et al., 2020; Elayyan, 2021; Ng & Tsang, 2021)) were conducted on the research fields as the struggle to result in students whom ready to face industrial revolution 4.0. To better adapting IR 4.0, teachers also seek the strategies to inculcate higher order thinking skills such as critical thinking, analytical skills, and other skills(Rosidin et al., 2019), (Gambari et al., 2015) Moreover, the integration of study can be perceived as a must for the current (Chawla et al., 2021; Hillmayr et al., 2020).

The analysis of studies still depends on some statistical analysis comprising Analysis of Covaiance (ANCOVA) with pre-test as covariate, independent t-test, qualitative analysis etc (Halimah et al., 2019), (Wahono et al., 2020). The prevailing analysis does not give a richer data of detailed results of students by considering item level. They only focus on reveals the data as whole instrument which neglecting potentials difficulties of students in details. Another issue is the use of raw score, indeed, do not meet the criteria to be treated as interval data (Saidfudin et al., 2010) to carry parametric statistics such as ANCOVA, t-test etc.

In the current study, we highlight the use of Rasch theory to analyze the effect of double-loop problem solving towards student's critical thinking. The proponent of Rasch model measurement are two theorems: 1) a more capable person has a higher probability of correctly responding to all the items provided. 2). An easier item is more likely to be answered correctly by all respondents or test-takers(Bond & Fox, 2015; Sumintono & Widhiarso, 2015). The tendency of current studies in the field of finding best teaching strategies pivots their analysis based on student's performance, while Rasch model emphasize their analysis for both person and items. In the field of science education, a plethora of studies utilize Rasch model only to validate instrument or they just utilize the theory for enacting linear data(Qudratuddarsi et al., 2019; Sadhu & Laksono, 2018). Some preceding studies (e.g., (Ibrahim et al., 2012; Ismail et al., 2017)) which have shown that the analysis of

learning outcomes using Rasch model yield a richer data to gauge student's understanding, however they apply the analysis for single test while this study has pre-test and post-test. Therefore, this study depicted the analysis using Rasch model which pivots on analysis on person performance and items.

METHOD

This study was a quasi-experimental study with a non-equivalent design or also called an untreated control-group design (Creswell, 2014). There were 59 pupils take part in the study and they were divided into two groups i.e., 28 students (males, females) control group and 31 students (males, females) in the experimental group. Both groups experienced pre-test before learning process and post-test after the intervention. While students in control class have learnt physics as daily practices, experiment class studied physics by applying the steps for double-loop problem solving method. In general, there are five stages of learning process comprising the identification phase, causal detection, tentative solutions, consideration of solutions and causal analysis.

Before running the study, as the first step, Physics Critical Thinking Skills (PCTS) instrument was formulated and validated. The instrument has 10 items with 5 indicators namely interpretation (Int), analysis (A), evaluation (E), inference (Inf) and explanation (Ex) as formulated based on (Facione, 2011). Each indicator consisted of two essay questions with scoring methods ranging from 0-4 points. PCTS instrument was pilot tested to ensure it to be valid and reliable instrument to measure the constructs. Winstep version 3.73 was utilized to analyze the data and the result of the analysis is depicted. According to DeVellis (2012), minimally acceptable reliability for person and item are 0.65, unexpectedly, person reliability is merely 0.64. It elicits that the respondents do not answer the PCTS consistently. Scientific instruments sometimes had low reliability score such as the some listed studies (e.g., (Hoe & Subramaniam, 2016; Yan & Subramaniam, 2018)). Number of participants can also be contributing factor of the low score. However, item reliability and Cronbach's alpha are in acceptable range, accounting for 0.95 and 0.66 respectively. The satisfying score indicates that the item has stability of score and elicits the positive and good correlation between a score of an individual item in the test and the total gained score for all items (Chua, 2013; Sumintono & Widhiarso, 2015). The next value to consider is separation either person separation index (1.34) or item separation index (4.21). Person separation index and item separation index estimate how well the questionnaire can distinguish between "person abilities" in terms of the latent trait, and how widespread the items are in defining both the easy and difficult items (Boone et al., 2014; Chan et al., 2014). Based on (Sumintono & Widhiarso, 2015), one equation to estimate separation (H) ={(4 x separation) + 1}/3. Based on the calculation, we found that person separation is 2.02, meaning that students can be agglomerated into two groups, while item can be separated into six groups. The bigger separation the better spread of either person or item (Adams et al., 2018).

Table 1. Validity and reliability of instrument

	Person	**Item**
Separation	1.34	4.21
Reliability	0.64	0.95
Cronbach′s alpha	0.66	

As the proof of construct validation, analysis of items in the current study utilized Rasch model. There are some fit statistics to measure such as mean square (MNSQ), tolerated Z-Standard (ZSTD) and Correlation Points (Pt Mea Corr). According to Boone, Staver, and Yale (2014), the criteria: (a) the value of accepted infit and outfit mean square (MNSQ): 0.5 <MNSQ <1,5 (b) the value of tolerated infit and outfit Z-Standard (ZSTD): -2.0 <ZSTD <+2,0 (c) the value of accepted Correlation Points (Pt Mean Corr) must be positive value. An item will be eliminated if the three indicator requirements cannot be fulfilled. Most of the items completely full the requisite score, but some items have a large score. The value suggests that the item fails to differentiate among students in terms of the targeted ability and thus may measure a different construct from the rest of the items. Obviously, a large outfit statistic is more problematic than a small outfit statistic (Liu et al., 2011).

Table 2. Item fit statistics

Item	Infit		Outfit		Pt Mea Corr
	MNSQ	ZSTD	MNSQ	ZSTD	
1	1.11	0.6	1.66*	2.4*	-0.03*
2	0.61	-2.0	0.70	-1.5	0.55
3	1.76*	2.7*	1.62*	2.0	0.38
4	1.54*	2.3*	1.52*	2.2	0.79
5	0.54	-2.3*	0.52	-2.2	0.67
6	0.97	0.1	0.55	-0.5	0.33
7	0.74	-0.8	0.64	-0.8	0.38
8	0.84	-0.6	0.73	-0.7	0.49
9	1.13	0.7	1.09	0.5	0.66
10	0.62	-1.8	0.69	-1.1	0.53
Mean	0.99	-0.1	1.00	-0.3	
SD	0.39	1.6	0.29	3.8	

Note: * item with a fit value outside acceptable score

The last analysis to consider is unidimensionality, a term to explains how much HMCTS instrument can explain the score as the proof of construct validation. The information is vital to identify the measured dimension or domain (Sick, 2010; Stone & Yeh, 2006). The explained variance of HMCTS instrument is 52.8%, where raw variance explained by persons and items are 8.5% and 44.3%.

Table 3. *Unidimensionality result*

Variance in Eigenvalue units	Eig	Obs (%)	Exp (%)
Raw variance explained by measures	11.2	42.8%	53.2%
Raw Variance explained by persons	1.8	8.5%	22.3%
Raw Variance explained by items	9.4	44.3%	24.0%
Raw unexplained variance (total)	10.0	47.2%	68.2%

RESULT AND DISCUSSION

Results of analysis the effects of double-loop problem solving teaching methods towards critical thinking skills methods using Rasch model are classified into student's abilities based on person measure and item analysis.

1.1 Student Critical Thinking Ability

In the beginning, both pre-test and post-test student's raw scores are converted to logit score to create interval data as the requirement of statistical analysis (Qudratuddarsi et al., 2019). The conversion result is visualized in a wright map, a map to present the relationship between item and respondent at single picture (Mi et al., 2021) as depicted in Figure 1. In the figure, the left slide is item location, where the alphabet code refers to Int (interpretation), A (analysis), evaluation (Ev), inference (Inf) and explanation (Ex), while the number is item order. In the right side, it is person spread, where the code refers to after (A) or before (B) intervention, control (C) or experiment (E) and person order. For instance, AC21 means post-test score (code A), control class (code C), student number 21.

Based on the map, we can see item A2 and item Int2 probably cannot be answered correctly by students in pre-test because person logit of all students below the item logit. Moreover, only students in experiment class have a good tendency to solve all the questions there. A2 question "A point of matter performs simple harmonic vibrations with the equation of displacement. When its kinetic energy is equal to 3 times its potential energy, then what is the deviation?" and Int2 "A simple pendulum has period *T* and string length *l*. So that the period becomes ½ T. How does the length of the rope change?" Both questions are the same type which directly ask mathematical equation and there is no practical illustration in the questions. To be able answer both questions correctly, students are required to do a complicated calculation. They have to remember mathematical formula such as energy kinetic, potential, and mechanic equation for item A2 and period equation at pendulum for item Int2. Their numeracy skills are measured in these questions such as multiplication and division without calculator, multiplication of practional number, trigonometry, fraction simplification, and square roots problems.

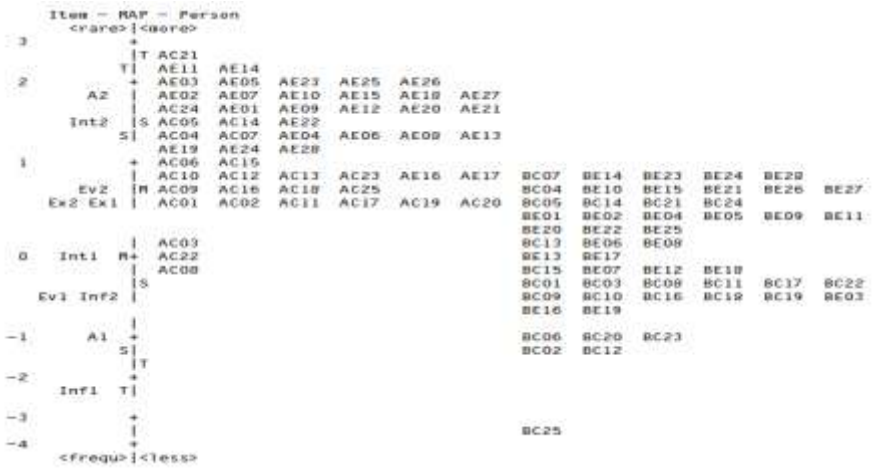

Figure 1. Wright Map of Item and Person Measure

However, items A1 and Inf1 are possibly the easiest items which can be answered correctly by students in pre-test and post-test. It is interesting to note that three items Ev2, Ex1, and Ex1 can be responded correctly by students after intervention, but they are very difficult for them before intervention. Item A1 "A springbed can move up and down past its equilibrium point if a load is applied. Why does this happen and what forces work?", and Inf1 "A spring is placed on a support post vertically. Then loads with different masses are alternately hung on the spring. Then, the students are given data about mass and 10T, T and T^2. They are requested to take conclusion". Both questions require pupils to apply their conceptual knowledge on giving scientific view to daily phenomenon and intrepet data. They just need to remember and memorize their lessons to be able to answer both questions and seems reasonable to be considered as the easiest questions.

Logit score of students, then, are analyzed using one-way Analysis of Covariance (ANCOVA) utilizing SPSS version 25. Based on the analysis, it was found that double-loop problem solving improve students critical thinking skills, F (1,52)= 19.026, P<0.001, the effect size for treatment = 0.276. The results obtained are in accordance with research conducted by (Roliyani, 2016) which states that the two-round problem-solving model can improve student learning outcomes and improve the quality of learning. In general, students in conventional class have less courage to actively involve in the class due to teacher domination. Similar results were shown in previous studies by (Furoidah et al., 2017; Sari et al., 2017) that conventional learning in schools using the direct instruction model results in students tending to listen, taking notes which causes students to feel bored in learning so that student physics learning outcomes are low. It is totally dfferent from double-loop problem problem which has three rounds namely identification phase, causal detection and tentative solutions that causes students tend to be active in the class. Their involvement yield a better critical thinking score. Phases of learning using the double loop problem-solving model can be used to train students' critical thinking abilities in physics. This is because students in each cycle of the learning phase show high curiosity to find solutions to problems presented by the teacher(Halimah et al., 2019).

Table 4. One-way ANCOVA results

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	11.176[a]	2	5.588	30.214	.000	.547
Intercept	73.801	1	73.801	399.029	.000	.889
Pre_test	2.037	1	2.037	11.015	.002	.181
Group	3.519	1	3.519	19.026	.000	.276
Error	9.248	50	.185			
Total	95.358	53				
Corrected Total	20.424	52				

R Squared = .547 (Adjusted R Squared = .529)

1.2 Item Analysis based on Rasch model

This analysis can be considered as novel, detailed and deeper analysis of student's critical thinking based on their performance on pre-test and post-test. In Rasch model, this analysis is part of Differential Item Functioning (DIF) by referring to significance level of Welch test, Mantel-Haenszel test and Chi-Square test. We consider the probability score (must lower than 0.05) to classify an item has shown significant different of score between control class and experiment class. The resut of the analysis is depicted in Figure 2, where (a) pre-test comparison between control and experiment class, (b) post-test comparison between control and experiment class.

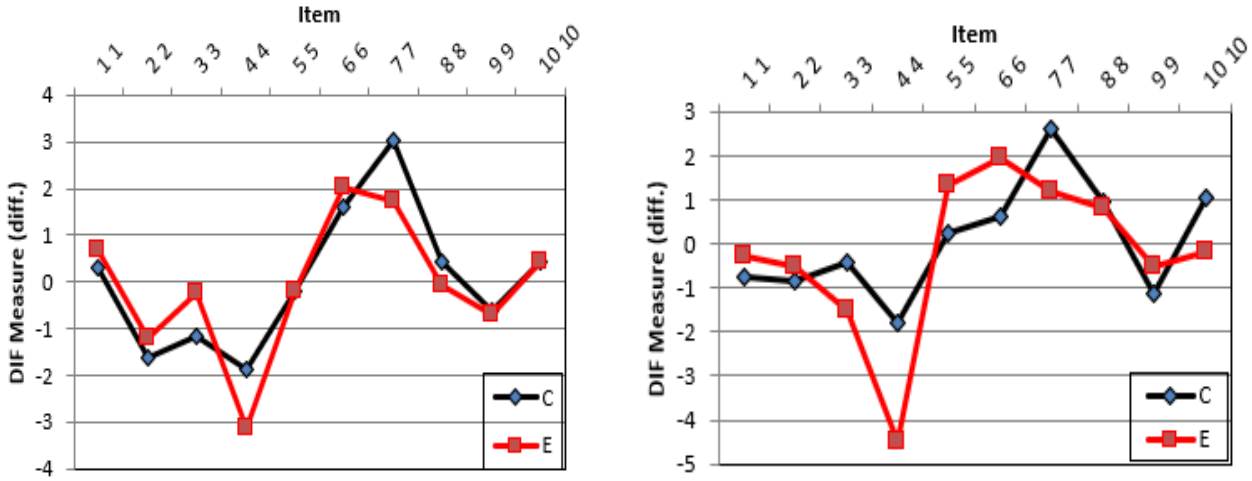

Figure 2. DIF Measure pre-test (left) and post-test (right)

Before intervention, there was no statistically significant different of scores at 70% items, and remained 40% after the intervention. The items are item Int1 "one application of harmonic vibrations in everyday life is on a pendulum clock that moves periodically. a) Why is the pendulum clock said to move periodically? b) Create a picture analysis of the force acting on the pendulum clock! c) What factors affect the movement of the pendulum clock?", item A1"A springbed can move up and down past its equilibrium point if a load is applied. Why does this happen and what forces work?", item Ev2 "(Figure 3.b) If the pendulum C is swung, which pendulum will swing with it? Why does it happen?" and item Inf2 "One application of harmonic vibrations is a pendulum clock. What will happen to the pendulum clock if it is taken into space (value of acceleration due to gravity)?". The first reason for this phenomenon is possibly because all of the questions can be understood quickly after reading the materials without further explanations from the teachers. In some point, the students are benefits from preceding lessons including vibration and force in their junior high school.

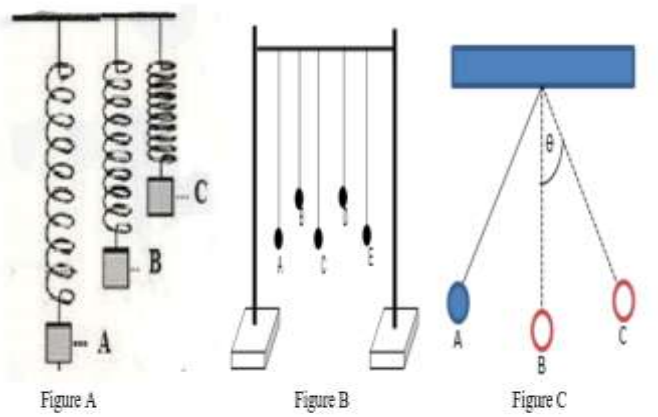

Figure 3. Question in item Ev2

Based on the test, it was found that 30% of items experience any difference and it uplifted to be 60% after the intervention. We divided items into 2 groups comprising 1) control group do better, 2) experiment group do better. In the first group, the items are item Ex1 "we often encounter harmonic motion events in life, one of which is a children's swing. If the rope on the swing can be shortened and lengthened. How is the rope swing if the child wants to quickly return to its original position?" and Int2 "A simple pendulum has period *T* and string length *l*. So that the period becomes ½ T. How does the length of the rope change?". This result is evident because students in control class, possibly, are familiar with the teaching methods and they feel convenience to get detailed explanations how to solve the problems. They also can focus on teacher's explanation regarding difficult conception in physics lessons, moreover they also get benefits from the questions on their worksheets. Based on researcher experience, the emphasis of the concepts also contributes to the results.

In the second group, the items are item Ev1 "(Figure 3.a) If you want to deviate some of the above springs that have different elasticity. What is the deviation if you want to get a large vibration period?", Inf1 "A spring is placed on a support post vertically. Then loads with different masses are alternately hung on the spring. Then, the students are given data about mass and 10T, T and T^2. They are requested to take conclusion", A2 "A point of matter performs simple harmonic vibrations with the equation of displacement. When its kinetic energy is equal to 3 times its potential energy, then what is the deviation?" and Ex2 "(Figure 3.c) At what point is the pendulum's kinetic and potential energies maximum? Why does it happen!". Majority of the questions is related to practical life which are discussed in details in experiment class, therefore they can get a better score. The students also construct their knowledge, they do not depend on teacher explanations. The results are also influence of conducting experiment where the students determine period in their worksheet.
Results of analysis the effects of double-loop problem solving teaching methods towards critical thinking skills mehods using Rasch model are classified into student's abilities based on person measure and item analysis.

CONCLUSION AND SUGGESTION

A. Conclusion

Based on logit's comparison as visualized by wright map and ANCOVA result, double-loop problem solving teaching methods improve student's critical thinking ability. This result is also supported by Wright-map which revealed the relationship between item and respondent at single picture. In the map, it is clearly can be seen that experiment group have higher logit compared to control group. Based on differential item functioning (DIF) analysis in Rasch model, analysis on each item revealed that no significant difference in some items, moreover some items are more preferable to control groups. Therefore, it can be said that Rasch model can give a better analysis of student's performance and deeper analysis on item analysis on the research of finding best teaching strategy.

B. Suggestion

Based on the conclusion that the Rasch model provides a better analysis of students' performance and deeper insights into item analysis for identifying effective teaching strategies, the following research suggestions can be made: Conduct comparative studies between the Rasch model and other psychometric models, such as Classical Test Theory (CTT) or Item Response Theory (IRT), Utilize the Rasch model to develop diagnostic frameworks that can identify specific areas where students struggle, providing targeted recommendations for improving teaching strategies and Apply the Rasch model in longitudinal studies to track changes in student performance over time, assessing the sustained impact of different teaching strategies.

ACKNOWLEDGMENTS

We would like to express our deepest gratitude to science education department, West Sulawesi University and Science and Technology Education Department, Institute for Innovative Learning, Mahidol University for their invaluable support and contributions to this research. We also sincerely thank the editorial team and reviewers of JPS for their valuable feedback and suggestions, which have significantly contributed to the improvement of this manuscript. We are grateful for the opportunity to share our work through this esteemed publication

REFERENCES

Adams, D., Sumintono, B., Mohamed, A., & Noor, N. S. M. (2018). E-learning readiness among students of diverse backgrounds in a leading Malaysian higher education institution. *Malaysian Journal of Learning and Instruction*, *15*(2), 227–256.

Apkarian, N., Henderson, C., Stains, M., Raker, J., Johnson, E., & Dancy, M. (2021). What really impacts the use of active learning in undergraduate STEM education ? Results from a national survey of chemistry , mathematics, and physics instructors. *PLoS ONE*, 1–15. https://doi.org/10.1371/journal.pone.0247544

Atmojo, S. E., Muhtarom, T., & Lukitoaji, B. D. (2020). The level of self-regulated learning and self-awareness in science learning in the covid-19 pandemic era. *Jurnal Pendidikan IPA Indonesia*, *9*(4), 512–520. https://doi.org/10.15294/jpii.v9i4.25544

Bogusevschi, D., Muntean, C. H., & Muntean, G. (2020). Teaching and Learning Physics using 3D Virtual Learning Environment : A Case Study of Combined Virtual Reality and Virtual Laboratory in Secondary. *Journal of Computers in Mathematics and Science Teaching*, *39*(1), 5–18.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model fundamental measurement in the human sciences* (3rd ed.). Routledge Taylor & Francis Group.

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer Netherlands.

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*, 133–139. https://doi.org/10.1016/j.sbspro.2014.03.658

Chawla, S., Tomar, P., & Gambhir, S. (2021). Design and implementation of IoT based Low cost , effective learning mechanism for empowering STEM education in India. *Turkish Journal of Computer and Mathematics Education*, *12*(13), 125–133.

Chua, Y. P. (2013). *Mastering research statistics*. McGraw-Hill Education.

Creswell, J. W. (2014). *Research design: qualitative, quantitative and mixed methods approaches* (4th ed.). Sage Publications.

Elayyan, S. (2021). The future of education according to the fourth industrial revolution. *Journal of Educational Technology & Online Learning*, *4*(1), 23–30.

Facione, P. A. (2011). *Critical Thinking: What it is and Why it Counts. Milbrae: California Academic Press.* California Academic Press.

Furoidah, A. Z., Indrawati, I., & Subki, S. (2017). Implementasi Model Discovery Learning Disertai Lembar Kerja Siswa dalam Pembelajaran Fisika Siswa di SMA. *Jurnal Pembelajaran Fisika*, *6*(3), 285–291.

Gambari, A. G., Yusuf, M. O., & Thomas, D. A. (2015). Effects of Computer-Assisted STAD, LTM and ICI Cooperative Learning Strategies on Nigerian Secondary School Students' Achievement, Gender and Motivation in Physics. *Journal of Education and Practice*, *6*(19), 16–28.

Ha, C. T., Thao, T. T. P., Trung, N. T., Huong, L. T. T., Dinh, N. Van, & Trung, T. (2020). A Bibliometric Review of Research on STEM Education in ASEAN: Science Mapping the Literature in Scopus Database, 2000 to 2019. *Eurasia Journal of Mathematics, Science and Technology Education*, *16*(10). https://doi.org/10.29333/ejmste/8500

Halimah, H., Sutrio, S., Nyoman, N., & Putu, S. (2019). The effect of double-loop problem-solving models on students' critical thinking ability. *Prisma Sains: Jurnal Pengkajian Ilmu Dan Pembelajaran Matematika Dan IPA*, *7*(2), 160–168.

Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S. I., & Reiss, K. M. (2020). Computers & Education The potential of digital tools to enhance mathematics and science learning in secondary schools : A context-specific meta-analysis. *Computers & Education*, *153*(April), 103897. https://doi.org/10.1016/j.compedu.2020.103897

Hoe, K. Y., & Subramaniam, R. (2016). On the prevalence of alternative conceptions on acid-base chemistry among secondary students: Insights from cognitive and confidence measures. *Chemistry Education Research and Practice*, *17*(2), 263–282. https://doi.org/10.1039/c5rp00146c

Ibrahim, F. M., Zubairi, Y. Z., Aziz, A. A., & Zaharim, A. (2012). Using the Rasch model to assess examination Beyond students' scores. *IEEE Global Engineering Education Conference (EDUCON)*, 1–5.

Ismail, A., Roslan, L., & Adnan, A. N. (2017). Assessment on course outcome performance using Rasch measurement model. *IEEE 9th International Conference on Engineering Education (ICEED)*, 110–113.

Liu, O. L., Lee, H., Linn, M. C., & Liu, O. L. (2011). An investigation of explanation multiple-choice items in science assessment. *Educational Assessment*, *16*(3), 164–184. https://doi.org/10.1080/10627197.2011.611702

Mi, J., Aksu, B., & Groote, S. L. De. (2021). Library and Information Science Research Validating students ' library experience survey using rasch model. *Library and Information Science Research*, *43*(1), 101071. https://doi.org/10.1016/j.lisr.2021.101071

Ng, O., & Tsang, W. K. (2021). Constructionist Learning in School Mathematics : Implications for Education in the Fourth Industrial Revolution. *CNU Review of Education*. https://doi.org/10.1177/2096531120978414

Qudratuddarsi, H., Sathasivam, R. V, & Hutkemri, H. (2019). Difficulties and correlation between phenomenon and reasoning tier of multiple-choice questions : A survey study. *Indonesian Research Journal in Education*, *3*(2), 249–264.

Roliyani, R. (2016). Upaya meningkatkan hasil belajar siswa melalui penggunaan model pembelajaran double loop problem solving. *Jurnal Pena Edukasi*, *3*(6), 560–566.

Rosidin, U., Suyatna, A., & Abdurrahman, A. (2019). Implementation of e-learning based-STEM on quantum physics subject to student HOTS ability. *Journal for the Education of Gifted Young*, *7*(2), 435–448.

Sadhu, S., & Laksono, E. W. (2018). Development and validation of an integrated assessment for measuring critical thinking and chemical literacy in chemical equilibrium. *International Journal of Instruction*, *11*(3), 557–572. https://doi.org/10.12973/iji.2018.11338a

Saidfudin, M., Azrilah, A., Rodzo'an, N., Omar, M., Zaharim, a, & Basri, H. (2010). Easier learning outcomes analysis using Rasch model in engineering education research. *Latest Trends on Engineering Education*, 442–447.

Sari, E. R., Pasaribu, M., & Sahena, S. (2017). Sari, E.R., Pasaribu, M., & Sahena, S. (2017). Pengaruh Model Discovery Learning terhadap Hasil Belajar Fisika pada Pokok Bahasan Kalor di SMP Negeri 2 Pamona Timur. Jurnal Inovasi dan Pembelajaran Fisika, 4(2), 119-126. *Jurnal Inovasi Dan Pembelajaran Fisika*, *4*(2), 119–126.

Sick, J. (2010). Rasch measurement in language education Part 5: Assumptions and requirements of Rasch measurement. *JALT Testing & Evaluation SIG Newsletter.*, *14*(2), 23–29.

Stone, C. A., & Yeh, C.-C. (2006). Assessing the dimensionality and factor structure of multiple-choice exams an empirical comparison of methods using the multistate bar examination. *Educational and Psychological Measurement*, 193–214. https://doi.org/10.1177/0013164405282483

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan Rasch pada assessment pendidikan*. Trim Komunikata.

Wahono, B., Lin, P., & Chang, C. (2020). Evidence of STEM enactment effectiveness in Asian student learning outcomes. *International Journal of STEM Education*, *1*, 1–18.

Yan, Y. K., & Subramaniam, R. (2018). Using a multi-tier diagnostic test to explore the nature of students' alternative conceptions on reaction. *Chemistry Education Research and Practice*, *19*, 213–226. https://doi.org/10.1039/c7rp00143f