



DEVELOPMENT OF PHYSIC PROBLEMS FOR THE FINAL ASSESSMENT IN HIGH SCHOOL LEVEL WITH RASCH MODEL ANALYSIS

S. Maulana^{1,✉}, Y. Supriyati¹, I M. Astra¹

¹University of Negeri Jakarta, Indonesia

Article Info

Article History:

Received April 2019

Accepted July 2019

Published Februari 2020

Keywords:

Physics Problem, Final Assessment, Rasch Model.

Abstract

The aim of this research is to develop Physics problems that are appropriate to be used for the final assessment at High School level with Rasch Model Analysis. Research method used in this study is Research and Development with the ADDIE model. Physics problems developed refers to the USBN blueprint issued by BSNP. Problems that have been developed will be validated by Physics experts, then tested on students, and analyzed for their quality using Rasch Model analysis. Validation results from Physics experts gave a score of 86.25% with very good criteria, it shows that Physics problems are feasible in terms of Physics material. The problems that have gone through the validation stage are then tested to several students and analyzed using the Rasch Model. Results of the analysis show that problems that have been developed have good validity and reliability. Problems that have been developed also have varying degrees of difficulty. Results of this analysis also show that problems that have been developed are not biased for a certain group. Based on the results, it can be concluded that Physics problems that have been developed are appropriate to be used for the final assessment at High School level.

©2020 Universitas Negeri Semarang
p-ISSN 2252-6617
e-ISSN 2252-6232

✉ Corresponding author:

S. Maulana

¹Universitas Negeri Jakarta, DKI Jakarta 13220

E-mail: slametmaulana1996@gmail.com

INTRODUCTION

Physics is one of fundamental knowledges which discuss about life phenomenon and can be used to predict a certain event by manipulating variable/s (Hırça, 2013; Hodosyová et al., 2015; Suyidno et al., 2018). Physics is one part of science that integrates with other science fields such as Mathematics, Chemistry, and Biology. In Indonesia, at the level of High School Education, Physics is one of the vocational subjects for the students who major in Mathematics and Natural Sciences.

At the end of a learning process, students will take several assessment activities to evaluate their knowledge that has been absorbed (Maba, 2017; Tighe-mooney et al., 2016). The assessment in Indonesia has been regulated by Minister of Education and Culture Regulation No. 23 year 2016 (Maulana & Ningtiyas, 2019). There're three types of assessment, one of them is an assessment by the education unit level. In Indonesia, the assessment is known as the National Standardized School Examination (Ujian Sekolah Berstandar Nasional or USBN).

Even though it has the word "National" in its name, the implementation of USBN is not same as the implementation of the National Examination. The National Examination is a test that has been organized nationally, so the problems that used in one school with the other schools must be the same, while the USBN is not the case. The results of observations show that problems used for the implementation of USBN are not from the national government. Problems used for USBN are divided into two, about 20% -25% of the problems are developed in Teacher Subject Deliberations (Musyawarah Guru Mata Pelajaran or MGMP) in certain cities / districts, while the remaining 75% - 80% are problems developed by each teacher in their respective schools.

Based on the observation, problems that developed by each school's teacher and used in the previous USBN are not in good quality. It is because of problems which have been developed are taken from irresponsible sources such as blog. This raises a big question about "Are problems used in the past USBN really able to measure what should be measured?". Therefore, the development of Physics problems for the implementation of USBN is needed.

Rasch Model is one of analytical techniques that not only can determine the quality of the item, but also knows the quality of test participants (Ishak et al., 2018). Rasch Model is an analysis technique of

one logic parameter that refers to difficulty level of an item in determining the quality of respondents (students) (Motta et al., 2015; Sasmoko et al., 2018). This is clearly different from analysis technique in general, which makes the number of questions answered correctly as a criteria for determining the quality of students.

The Rasch Model provides several advantages (Simpelaere et al., 2017). In addition to knowing whether the problems that have been developed are valid and reliable (Czuba et al., 2016; Ee et al., 2018; Hassan et al., 2017), as well as the level of difficulty of each problems, Rasch Model can also provide information about the quality of test instrument as a whole (Susongko, 2016), the quality of students, and interactions that occur between tests used with students (Cecilio-fernandes et al., 2017; Kean, Bisson et al., 2018; Lang et al., 2019; Morán et al., 2018; Wang et al., 2017). Quality of respondents (students) can also be obtained better (students with low, medium, and high ability), because the quality of students is seen from questions answered with a certain level of difficulty, not from the number of questions answered correctly (Lo et al., 2015; Mamat et al., 2018; Zubairi & Kassim, 2006). Rasch Model is also capable of analyzing problems that have the potential to be biased towards certain groups, thus helping educators and problem developers to consider improving a problem item later (Baghaei et al., 2017; Ismail et al., 2015; McLaughlin et al., 2016; Sumintono, 2018). Rasch Model is also useful for detecting the presence of students with inappropriate response patterns (Imran et al., 2017; Ireland et al., 2018; Li et al., 2016). The pattern of responses that are not appropriate is the pattern of answers given with their incompatibility (Maat & Rosli, 2016; Mohammad et al., 2016). This is used to indicate the consistency of students' thinking to know students who are careless or lucky and students who cheat in test (Saidi & Siew, 2019).

Based on the background issue which have been described above, the aim of this research is to develop Physics problems that are appropriate to be used for the final assessment at High School level with Rasch Model Analysis.

METHODS

This research was conducted from January 2019 to April 2019 in the Research and Development of Physics Education Laboratory, Universitas Negeri Jakarta. The trial test of problems that have been developed was carried out in several high schools in Bekasi city, with 100 students as research subject. This research is research and development with the ADDIE model.

ADDIE's research and development model consists of five core steps, they are Analysis, Design, Development, Implementation, and Evaluation (Muruganantham, 2015). In analysis step, several activities were carried out, those are observing the current USBN implementation, analyzing the problems that had been used in the USBN a few years before, and conducting literature studies on USBN blueprint issued by BSNP. In design step, a framework and design of Physics problems will be designed to be developed, creating a blueprint of problems which will be developed, along with the problem rubric for the description type. In development step, the development of Physics problems for the implementation of USBN is carried out, where the questions that have been developed will be analyzed by Physics experts to assess their quality. Input from experts was used as material to revise the problems before being tested in the field. In implementation step, questions that have been declared feasible in terms of Physics material are tried out to several students. Results of the trials were then analyzed using Rasch Model analysis to obtain information about the quality of instrument as a whole, the quality of problems that had been developed, and the quality of students used in conducting the trial. Finally, in evaluation step, a conclusion is made whether the questions that have been developed are suitable to be used for USBN Physics material or not.

In this study, validation data from Physics experts were obtained using a validated questionnaire beforehand. The questionnaire used was Likert scale 1-4. Validation results from Physics experts were processed and analyzed to determine the feasibility of the questions that had been developed in terms of the material. While the results of trials by a number of students were processed and analyzed by Rasch Model analysis with the help of the Winsteps application.

RESULTS AND DISCUSSION

Product

In this study, a research product is Physics problems that will be functioned for the implementation of USBN in High School level. The problems that have been developed refer to the USBN blueprint issued by BSNP.



Figure 1. Display of one of the problems that have been developed

Theoretical Validation

Problems that have been developed then validated by two experts in terms of Physics material. Theoretical validation by Physics experts aims to determine the quality of the questions that have been developed from Physics material aspects. The results of theoretical validation by material experts are shown in the following graph.

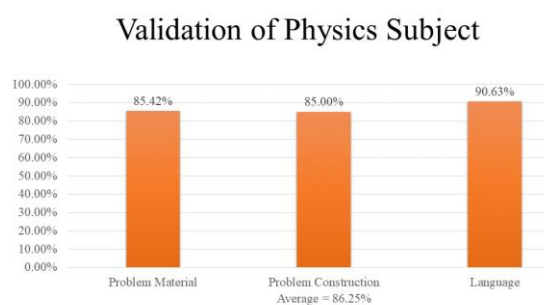


Figure 2. Theoretical validation from Physics experts

The graph above shows that the quality of problems that have been developed has very good criteria with an average score of 86.25%. The more complete results of the material expert validation in the "material problems" section showed an average score of 85.42%, the "problem

construction" section was 85.00%, and the "language" section was 90.63%, where every section has very good criteria. In courtesy, the results of this validation show that the quality of the Physics problems that have been developed have very decent quality to be used on USBN in High School level for Physics subject.

Field Test and Analysis

The trial was carried out by asking students to answer Physics problems that had been arranged in one complete instrument. Answers given by students are then processed and analyzed using the Rasch Model. The analysis carried out is an analysis to determine the quality of the response pattern between respondents (students) with items (problems), overall quality of the instrument, overall quality of students, validity of the items, level of difficulty of the items, individual ability, and degree of conformity individuals in responding (answering questions).

Summary statistic shows overall info about the quality of the overall response pattern, overall quality of the instrument, and the interaction between person and item. The Summary statistic has an average value of -0.88 logit which shows the value of all respondents in working on the items given. The average value that is smaller than the logit value of 0.0 indicates the tendency of respondents to be smaller than the difficulty level of the items. In addition, person reliability and item reliability in data are used to determine the level of reliability of respondents and items. If it's less than 0.67, it means weak reliability; 0.67 to 0.80 is enough, 0.80 to 0.90 is good; 0.91 to 0.94 is very good; and more than 0.94 is said to have special reliability. Person reliability has a value of 0.56 where good consistency of answers from respondents has weak reliability, while item reliability has a value of 0.89, where the quality of the items in the first test device shows good reliability.

Item fit explains whether the item functions normally in measuring or not. If the problem is found not to be fit, it is an indication that there is a possibility of misconception in students for the item. This information will be used as a reference to improve the quality of the items which are not fit. The outfit means-square, outfit z-standard, and correlation measure point are the criteria used to see the level of item quality. The item is declared valid if it has one of the criteria, but if nothing is fulfilled, then it can be ascertained that the item is not good, so it needs to be repaired or replaced. The suitability criteria for the items: (1) The received Outfit Mean Square (MNSQ) values range from $0.5 < \text{MNSQ}$

< 1.5 ; (2) The value of Outfit Z-Standard (ZSTD) received ranges from $-2.0 < \text{ZSTD} < +2.0$; and (3) The value of Point Measure Correlation (Pt Mean Corr) ranges from $0.4 < \text{Pt Measure Corr} < 0.85$. In this research, it was found that each item fulfilled at least one of the criteria of an item fit. This shows that the problems that have been developed are valid and able to measure what should be measured.

The approach to use the Rasch Model also provides an interpretation of the difficulty level in item measure. Through the Winsteps application, the difficulty level of the problem can be seen from the logit value. A high logit value indicates a high level of difficulty, and vice versa. Positive logit values show difficult and negative items for easy items. While the standard deviation (SD) value can also be combined to classify the difficulty level. $0.00 \text{ logit} + 1\text{SD}$ is a group of difficult problems, more than $+1\text{SD}$ is a group of very difficult problems; $0.00 \text{ logit} - 1\text{SD}$ is a group of easy problems, and smaller than 1SD is a group of very easy problems. In the test instruments that have been developed, problems number 38, 11, 34, 40, 1, 2, and 20 show very difficult problems; problems number 39, 3, 25, 31, 12, 21, 32, 35, 10, 13, 36, 8, 19, and 14 including difficult problems; problems number 9, 24, 7, 18, 30, 37, 12, 22, 28, 5, and 33 including easy problems, as well as problems number 23, 27, 29, 6, 4, 16, 26, and 25 including the problems that very easy.

Besides for item analysis, the Rasch Model can also be used to analyze the respondent's ability (person measure). Through Rasch Model analysis, respondents can be divided into a several various levels of ability. The average logit value is used as the identifier of the respondent group. While the standard deviation (SD) value is combined to classify the respondent's ability level. In the range of "average logit - 1SD " to "average logit + 1SD " is a group of respondents with moderate ability, more than "average logit + 1SD " is a group of highly capable respondents, and smaller than "average logit average - 1SD " is a group of low-ability respondents. The results show that the average logit value is -0.88 and the SD value is 0.53. So that the group of highly capable respondents must have a logit value greater than -0.30; the group of respondents who are capable have logit values ranging from -1.46 to -0.30; and groups of respondents with low abilities have a logit value of less than -1.46.

Person Fit is used to detect the presence of individuals who have inappropriate response patterns. The different response patterns are patterns of incompatibility of answers given based on their ability. This is used to indicate the consistency of respondents' thinking and to find out respondents who are careless or lucky and respondents who cheat in test. The outfit means-square, outfit z-standard, and point measure correlation are the criteria used to see the level of individual fit (person fit). Individual suitability is declared valid if it fulfills one of the criteria, but if nothing is fulfilled, it is certain that the individual is inconsistent, careless, lucky, or cheating. Following are individual suitability criteria: (1) The value of Outfit Mean Square (MNSQ) received ranges from $0.5 < \text{MNSQ} < 1.5$; (2) The value of Outfit Z-Standard Outfit (ZSTD) received ranges from $-2.0 < \text{ZSTD} < +2.0$; and (3) The value of Point Measure Correlation (Pt Mean Corr) ranges from $0.4 < \text{Pt Measure Corr} < 0.85$. The results of person fit show that there were no respondents who were inconsistent, careless, lucky, or cheating. But when viewed on a scalogram, it appears that 97 respondents including respondents who are careless because the third easiest item (number 16) cannot answer it, but the hardest item (number 38) can be answered. In addition, respondents 11 included respondents who were lucky because they were able to answer the most difficult items (number 38).

Discussion

In this study, a research product is Physics problems that will be functioned for the implementation of USBN in High School level. The problems refer to the USBN blueprint issued by BSNP.

The problems that have been developed then validated by several experts in Physics material aspects. The results of the validation of material experts show that the quality of the problems that have been developed has very good criteria with an average score of 86.25%. The more complete results of the material expert validation in the "material problems" section showed an average score of 85.42%, the "problem construction" section was 85.00%, and the "language" section was 90.63%, where every section has very good criteria. In courtesy, the results of this validation show that the quality of the Physics problems that have been developed have very decent quality to be used on USBN in High School level for Physics subject.

Physics problems that have been developed and validated by several experts in terms of Physics material, then it will be tested in the field. The trial was carried out by asking for responses from several students by answering the problems that had been developed. Each given answer was processed and analyzed using the Rasch Model to find out not only the quality of the items developed, but also the quality of the test participants.

Summary statistic data shows that the Summary statistic has an average value of -0.88 logit which shows the value of all respondents in working on the items given. The average value that is smaller than the logit value of 0.0 indicates the tendency of the respondents to be smaller than the difficulty level of the items. In addition, person reliability and item reliability in data are used to determine the level of reliability of respondents and items. If it's less than 0.67, it means weak reliability; 0.67 to 0.80 is enough, 0.80 to 0.90 is good; 0.91 to 0.94 is very good; and more than 0.94 is said to have special reliability. Person reliability has a value of 0.56 where good consistency of answers from respondents has weak reliability, while item reliability has a value of 0.89, where the quality of the items in the first test device shows good reliability.

Item fit explains whether the item functions normally in measuring or not. If the problem is found not to be fit, it is an indication that there is a possibility of misconception in students of the item. This information will be used as a reference by researchers to improve the quality of the problems if items that are not fit are obtained. Based on the results of the study, it appears that each item at least fulfills one of the criteria of an item fit. This shows that the problems that have been developed are valid and able to measure what should be measured.

The approach to using the Rasch Model also provides an interpretation of the difficulty level of an item. Through the Winsteps application, the difficulty level of the problem can be seen from the logit value. The result show that problems number 38, 11, 34, 40, 1, 2, and 20 show very difficult problems; problems number 39, 3, 25, 31, 12, 21, 32, 35, 10, 13, 36, 8, 19, and 14 including difficult problems; problems number 9, 24, 7, 18, 30, 37, 12, 22, 28, 5, and 33 including easy problems, as well as problems number 23,

27, 29, 6, 4, 16, 26, and 25 including very easy problems.

Rasch Model can also be used to analyze respondents. Based on the data, the students with high ability among them are student number 17 and 23. The students with moderate ability include number 31 and 42, while students with low ability are number 13 and 19.

Person Fit is used to detect the presence of individuals who have inappropriate response patterns. The different response patterns are patterns of incompatibility of answers given based on their ability. This is used to indicate the consistency of respondents' thinking and to find out respondents who are careless or lucky and respondents who cheat in test. The results of person fit show that there were no respondents who were inconsistent, careless, lucky, or cheating. But when viewed on a scalogram, it appears that 97 respondents including respondents who are careless because the third easiest item (number 16) cannot answer it, but the hardest item (number 38) can be answered. In addition, respondents 11 included respondents who were lucky because they were able to answer the most difficult items (number 38).

The results of the analysis show that the quality of Physics problems that have been developed can be declared feasible to be used for the implementation of USBN for High School level. In addition to providing an analysis of the quality of the problems that have been developed, the implementation of Rasch Model also provides information on the quality of students who carry out the test. This informs educators in improving and / or improving the quality of their students. This is in line with some of the results of previous studies which explained that Rasch Model provides more in-depth information about the quality of students (Zamri & Nordin, 2015), thus helping educators in evaluating the learning process in the classroom (Mursidi & Soeharto, 2016; Rahmani, 2018; Suranata et al., 2018). In addition, the quality of the items (problems) that are informed can be more accountable because the quality of the questions is analyzed using the difficulty level of each item

CONCLUSION

Based on the results and discussion, it can be concluded that Physics problems that have been developed are appropriate to be used for the final assessment at High School level.

ACKNOWLEDGMENT

In this section, thanks are given to several parties who helped in carrying out this research. Thank you to Mr. Esmar Budi and Mr. Iwan Sugihartono as Physics material expert validators. In addition, we also give thanks to all education personnel / citizens of SMA Negeri 1 Bekasi, SMA Negeri 2 Bekasi, SMA Negeri 4 Bekasi, SMA Negeri 9 Bekasi, SMA Negeri 14 Bekasi, and SMA Negeri 20 Bekasi for providing convenience in carrying out this research trial test to several of their students.

REFERENCES

- Baghaei, P., Yanagida, T., & Heene, M. (2017). Development of a Descriptive Fit Statistic for the Rasch Model. *North American Journal of Psychology*. 19(1) 155–168.
- Cecilio-fernandes, D., Medema, H., Collares, C. F., Schuwirth, L., Cohen-schotanus, J., & Tio, R. A. (2017). Comparison of Formula and Number-right Scoring in Undergraduate Medical Training: a Rasch Model Analysis. *BMC Medical Education*. 17(192) 1–9.
- Czuba, K. J., Kersten, P., Kayes, N. M., Smith, G. A., Barker-collo, S., Taylor, W. J., & Mcpherson, K. M. (2016). Measuring Neurobehavioral Functioning in People With Traumatic Brain Injury: Rasch Analysis of Neurobehavioral Functioning Inventory. *Journal of Head Trauma Rehabilitation*. 31(4) 59–68.
- Ee, N. S., Yeo, K. J., & Kosnin, A. bt M. (2018). Item Analysis for the Adapted Motivation Scale Using Rasch Model. *International Journal of Evaluation and Research in Education*. 7(4) 264–269.
- Hassan, S. A., Fauzi, A., Ayub, M., & Bakar, K. A. (2017). Rasch Model Analysis for Malaysian Internet and Sexual Activities Inventory. *International Journal of Academic Research in Business and Social Sciences*. 7(4) 711–718.
- Hırça, N. (2013). The Influence of Hands on Physics Experiments on Scientific Process Skills According to Prospective Teachers'

- Experiences. *European Journal of Physics Education*. 4(1) 6–14.
- Hodosyová, M., Útla, J., Vnuková, P., & Lapitková, V. (2015). The Development of Science Process Skills in Physics Education. *Procedia - Social and Behavioral Sciences*. 186 982–989.
- Imran, M., Junaidi, M., Zain, S., Hamid, R., Ezlin, N., & Basri, A. (2017). Initial Impact Analysis of STEM Educational Model by Innovative Sustainable Home (IG-Home) using the Rasch Measurement Model. *Journal of Engineering Science and Technology*. 91–98.
- Ireland, M. J., Goh, H. E., & Marais, I. (2018). A Rasch Model Analysis of the Emotion Regulation Questionnaire. *Journal of Applied Measurement*. 19(3) 1–13.
- Ishak, A., Osman, M. R., Mahaiyadin, M. H., & Tumiran, M. A. (2018). Examining Unidimensionality of Psychometric Properties via Rasch Model. *International Journal of Civil Engineering and Technology*. 9(9) 1462–1467.
- Ismail, R., Noh, M. Z. M., Ismail, N. H., & Tamil, A. M. (2015). A Tale of Two Construct Validation Analysis: Rasch Model and Exploratory Factor Analysis Approach for Three-Factor Eating Questionnaire (TFEQ-R21) among Malaysian Male Workers. *The Medical Journal of Malaysia*. 70(3) 169–176.
- Kean, J., Bisson, E. F., Brodke, D. S., Biber, J., & Gross, P. H. (2018). An Introduction to Item Response Theory and Rasch Analysis: Application Using the Eating Assessment Tool (EAT-10). *Brain Impairment*. 19(1) 91–102. <https://doi.org/10.1017/BrImp.2017.31>
- Lang, W. S., Moore, L. L., & Wilkerson, D. J. (2019). *Assessment of Teacher Dispositions with the ETQ2: A Guided-Reflection and Rasch Model Analysis*.
- Li, Y., Root, J. C., Atkinson, T. M., & Ahles, T. A. (2016). Examining the Association between Patient-Reported Symptoms of Attention and Memory Dysfunction with Objective Cognitive Performance: A Latent Regression Rasch Model Approach. *Archives of Clinical Neuropsychology*. 31 365–377.
- Lo, C., Liang, W., Hang, L., Wu, T., Chang, Y., & Chang, C. (2015). A Psychometric Assessment of the St. George's Respiratory Questionnaire in Patients with COPD using Rasch Model Analysis. *Health and Quality of Life Outcomes*. 13(131) 1–16.
- Maat, S. M., & Rosli, M. K. (2016). The Rasch Model Analysis for Statistical Anxiety Rating Scale (STARS). *Creative Education* 7(1) 2820–2828.
- Maba, W. (2017). Teachers' Perception on the Implementation of the Assessment Process in 2013 Curriculum. *International Journal of Social Sciences and Humanities*. 1(2) 1–9.
- Mamat, M. N., Temyati, Z., Mahamood, S. F., & Musa, H. (2018). Simplified Reliable Online Essay Test Marking for Massive Open Online Course (MOOC) using Rasch Model Analysis. *Social and Management Research Journal*. 15(2) 23–34.
- Maulana, S., & Ningtiyas, H. A. (2019). The Implementation of Chapter II Article 3 in Government Education and Culture Regulation. In *Advances in Social Science, Education and Humanities Research*. 277 pp. 81–84.
- McLaughlin, J. E., Singer, D., & Cox, W. C. (2016). Candidate Evaluation Using Targeted Construct Assessment in the Multiple Mini-Interview: A Multifaceted Rasch Model Analysis. *Teaching and Learning in Medicine*. 0(0) 1–7.
- Mohammad, H., Hassan, F., Rashid, R. A., & Yaman, S. K. (2016). *Dimensionality Analysis of Technical Competency for Malaysian Construction Managers*. In UNIMAS STEM ENCON.
- Morán, V. E., Cupani, M., Azpilicueta, A. E., Piqueras, J. A., & Garcia-lopez, L. J. (2018). Rasch Model analysis of the Brief Version of the Social Phobia and Anxiety Inventory (SPAI- B) in Argentinean and Spanish samples. *Revista de Psicopatología y Psicología Clínica*. 23(3) 211–220.
- Motta, C. da, Carvalho, C. B., Castilho, P., & Pato, M. T. (2015). *Assessment of Emotional Management ability: a Rasch Model Analysis of the brief Situational Test of Emotional Management (STEM-B)*. IOP Conf. Series: Materials Science and Engineering.
- Mursidi, A., & Soeharto. (2016). An Introduction: Evaluation of Quality Assurance for Higher Educational Institutions using Rasch Model. *Journal of Education, Teaching and Learning*. 1(1) 1–6.

- Muruganantham, G. (2015). Developing of E-content package by using ADDIE model. *International Journal of Applied Research*. 1(3) 52–54.
- Rahmani, B. D. (2018). Differential Item Functional Analysis on Pedagogic and Content Knowledge (PCK) Questionnaire for Indonesian Teachers using RASCH Model. *Journal of Physics: Conference Series*.
- Saidi, S. S., & Siew, N. M. (2019). Reliability and Validity Analysis of Statistical Reasoning Test Survey Instrument using the Rasch Measurement Model. *International Electronic Journal of Mathematics Education*. 14(3) 535–546.
- Sasmoko, Abbas, B. S., Indianti, Y., & Widhoyoko, S. A. (2018). Indonesian Teacher Engagement Index: a Rasch Model Analysis. *IOP Conf. Series: Materials Science and Engineering*.
- Simpelaere, I. S., Nuffelen, G. Van, Bodt, M. De, Vanderwegen, J., & Hansen, T. (2017). Validation of the Dutch version of the Swallowing Quality-of-Life Questionnaire (DSWAL-QoL) and the adjusted DSWAL-QoL (aDSWAL-QoL) using item analysis with the Rasch model : a pilot study. *Health and Quality of Life Outcomes*. 15(66) 1–16.
- Sumintono, B. (2018). Rasch Model Measurements as Tools in Assessment for Learning. In *Advances in Social Science, Education and Humanities Research*. 173 pp. 38–42.
- Suranata, K., Rangka, I. B., Ifdil, I., Ardi, A., Susiani, K., Prasetyaningtyas, W. E., ... Rahim, R. (2018). Diagnosis of students zone proximal development on math design instruction : A Rasch analysis. *Journal of Physics: Conference Series*.
- Susongko, P. (2016). Validation of Science Achievement Test with the Rasch Model. *Jurnal Pendidikan IPA Indonesia*. 5(2) 268–277.
- Suyidno, Nur, M., Yuanita, L., Prahani, B. K., & Jatmiko, B. (2018). Effectiveness of Creative Responsibility Based Teaching (CRBT) Model on Basic Physics Learning to Increase Student's Scientific Creativity and Responsibility. *Journal of Baltic Science Education*. 17(1) 136–151.
- Tighe-mooney, S., Bracken, M., & Dignam, B. (2016). Peer Assessment as a Teaching and Learning Process : The Observations and Reflections of Three Facilitators on a First-Year Undergraduate Critical Skills Module. *All Ireland Journal of Teaching and Learning in Higher Education*. 8(2).
- Wang, W., Guedj, M., Bertrand, V., Fouquier, J., Jouve, E., Commenges, D., ... Attarian, S. (2017). A Rasch Analysis of the Charcot-Marie-Tooth Neuropathy Score (CMTNS) in a Cohort of Charcot-Marie-Tooth Type 1A Patients. *PLOS ONE*. 12(1) 1–14.
- Zamri, A., & Nordin. (2015). Modeling a Multiple Choice Mathematics Test with the Rasch Model. *Indian Journal of Science and Technology*. 8(12) 1–6.
- Zubairi, A. M., & Kassim, N. L. A. (2006). Classical And Rasch Analyses Of Dichotomously Scored Reading Comprehension Test Items. *Malaysian Journal of ELT Research*. 2 1–20.