The Rasch-rating scale model to identify learning difficulties of physics students based on self-regulation skills

Habibi Habibi[1], Jumadi Jumadi[2], Mundilarto Mundilarto[3]
[1]Doctor Candidate, Universitas Negeri Yogyakarta, Indonesia
[1]Department of Physics Education, IKIP Mataram (Undikma), Indonesia
[2,3]Department of Physics Education, Universitas Negeri Yogyakarta, Indonesia

ABSTRACT

This study applies the unidimensional Rasch Model assumptions in identifying physics learning difficulties based on students' self-regulation abilities. A total of 126 physics teacher candidates have been observed for one semester. Self-project as a learning strategy has been used. Data were collected using 20 items in rating scales and then analyzed quantitatively to get feasibility in measuring self-regulation skills. The results have shown that the profile items analyzed by the Rasch Model are feasible to measure self-regulation skills through a self-sustaining project strategy. Most physics teacher candidates have a medium ability of 51% in the process of self-regulation, high = 33%, and low = 16%. The implications of applying self-projects to the processes of self-regulation are discussed in this article.

Corresponding Author:

Habibi Habibi,
Department of Physics Education,
IKIP Mataram (Undikma), Indonesia
Pemuda Street No. 59A, Mataram, Postcode: 83125, Indonesia
Email: habibi27021983@yahoo.com

1. INTRODUCTION

Phenomena in everyday life that have been integrated in physical concepts can motivate learners to easily understand. Conscious or not, that all life has been surrounded by physical phenomena with all its complexity [1]. Although physics is very important to solve various problems in daily life, unfortunately many students consider it very complicated then disrupts their motivation and interest to learn it [2]. However, the complexity of cases in physics becomes a challenge for each learner to make it easy, interesting, and fun [3].

According to English [4] that motivation is closely related to the ability of self-regulation which has a positive impact on responsibilities and learning difficulties. This means that learning difficulties can be detected through self-regulation skills indicators. In addition, self-regulation is a psychological condition that can be detected using a psychometric scale test. Psychological conditions include perceptions, opinions, and attitudes of individuals or groups of people about various social phenomena [5-6].

A test must be categorized as valid to be used in interpreting the variables that are measured correctly [7]. Previous studies related to the psychology of cognitive processes have proven that self-regulation skills have been used to predict learning achievement [8-12]. However, the inability to regulate learning behavior is closely related to differences in levels of learning difficulties [13-14]. This allows that measurement items that have been used are not fit (invalid) with variables that have been measured.

This article describes the feasibility of the Rating Scale Model instrument to identify learning difficulties through the four strategies of self-regulation processes as follows: planning, controlling,

evaluating, and reinforcing. Self-regulation is a dimension of metacognitive knowledge used to regulate cognitive processes as a basis for planning learning strategies [15].

Each person has the characteristics to manage his cognitive processes which have implications for the application of learning strategies. If everyone has a diversity of self-regulation processes, then it is impossible to interpret [16-17]. This has contradicted the assumption of a unidimensional Rasch Model, [18]. Fischer [19] has asserted that the existence of different processes on the same continuum of abilities causes measurements not to be linear and impossible to compare. The implication requires the same initial assumptions about self-regulation strategies to easily generalize the diversity of individual learning difficulties found.

Rasch Rating Scale models (Rasch-RSM) are ordered categories that are used to view items as a representation of measured sample behavior. This means that the data collected is in the form of opinion scale or a latent attitude. The main assumption of RSM is that the assessment of response categories of items must be the same where the score must increase constantly according to the level of difficulty [20-21]. The Rasch model has better measurement accuracy than classical test theory (CTT). The Rasch model is able to process the evaluation of test results iteratively through calibration to find the optimal composition and meet the model criteria [22-24]. This advantage is not complemented by CTT; therefore it has the limitation of requiring more test items to get quality measurements.

The Rasch Model provides analysis to verify the assumptions of the items used. Rasch Modeling also provides estimates of the specific characteristics of the difficulty level of items in certain stages based on probability [25-28]. The probability approach accommodated aims to keep the Rasch Model not deterministic, therefore measuring objects can be identified more carefully. The probability of the Rasch Model for RSM has been developed based on the Partial Credit Model (PCM) probability equation as follows:

$$P_{nix} = \frac{exp\,(\beta_n - \delta_{ij})}{1 + exp\,(\beta_n - \delta_{ij})} \tag{1}$$

Where β is a component of the ability level, and δ_{ij} is the specific characteristic of the difficulty level of item i at each step j, whereas in RSM probability, the value δ_{ij} is described as $\delta_i - \tau_{ij}$ hence the equation becomes as follows:

$$P_{nix} = \frac{exp\,(\beta_n - \delta_i - \tau_{ij})}{1 + exp\,(\beta_n - \delta_i - \tau_{ij})}\,; \text{for } x = 1, 2, 3, ..., m_i \tag{2}$$

The probability equation of the Rasch Model shows that δ_i is the difficulty levels of response items, and τ_{ij} is the specific characteristics of the difficulty levels of response items i in each step of j. While items that have been tested are declared fit to be used if the INFIT T value is in the range of $\pm\,2$ while the standard error (α) is 5% [17].

The application of rating scale has been carried out in various social studies to detect a person's behavior such as: identifying benchmarks from the politomus scale [18], interpreting the strategies of student learning motivation [29], explaining understanding concepts from candidates' elementary school teacher, [30], and streamlining measurements to improve teacher social skills [31]. This has proven that the Rating scale model is flexible to be applied in various continuum involving the affective domain.

2. RESEARCH METHOD
2.1 Cases and instruments

The stability of the estimated results on the expected logit scale is needed. In specific cases using a scale of $\pm\,1$ logit (logarithm odds unit) to obtain the stability of the estimated data for a confidence level of 99% is required that the number of cases (sample size) is feasible at least for 50 people [32]. Therefore the sample of this study is 126 people who are prospective physics teachers in the early semester of 2018 in one of private university, Indonesia.

Teacher candidates who have been taken with purposive sampling are conditioned by project-based learning for one semester to facilitate self-regulation strategies. Through the application of this learning model, the ability of self-regulation of prospective teachers can be assumed to be homogeneous in each step of the project. The steps of the self-regulation strategy identified are as follows: planning, monitoring, evaluating, and reinforcing. Identification is done through the provision of questionnaires that have been adapted to self-regulation strategies. The questionnaire consisted of 20 items that used a four-point rating

scale namely: 1 = Completely Untrue, 2 = Untrue, 3 = True, 4 = Completely True. The questionnaire was given at the end of the semester to find out the learning strategies used while using project-based learning.

2.2 Data analysis

The probability of a Rasch Model is generated from the raw data analyzed using the QUEST application to find a fit model and estimate the teacher's self-regulation ability levels. The QUEST application provides a wider range of community measurement and research facilities based on the latest psychometric methods from heterogeneous item response models, multidimensional responses, latent regression models, and informs logical values. The QUEST program is freeware that can be downloaded at https://www.rasch.org/software.htm.

The central elements in the QUEST program are item response theories (IRTs) that have been adjusted to the Rasch Model (RM). The QUEST program uses unconditional (UCON) or a joint maximum likelihood to estimate parameter items on raw scores [33]. The raw score on a scale r will be converted to a logit scale that shows a person's ability (b) as (3):

$$b = \log [(r/ (L-r)] \tag{3}$$

For L is the number of activities (items). Whereas to find out the level of difficulty of item (d), however the value of r can be converted into a logit scale as in the equation as (4):

$$d = \log [(N-S)/S)] \tag{4}$$

Where N is the number of tests (case / person) and S is the score of the item [34]. The size of S for politomus' scaled measurement data in the QUEST program becomes w_{ij} and the difficulty level of d will be changed to δ_{ij}. The equation for RSM in the QUEST program is as (5):

$$P (X_{ni} = x_{ni}) \frac{exp \sum_{j=0}^{x_n} w_{ij} (\beta_n - \delta_i - \tau_{ij})}{\sum_{k=0}^{ni} exp \sum_{j=0}^{x_n} w_{ij} (\beta_n - \delta_i - \tau_{ij})} \tag{5}$$

where β_n is a component of the ability level of the test (case / person) n, w_{ij} is the score for step j on an item i, δ_i informs us of the difficulty levels of the items, and τ_{ij} is the specific characteristic of the difficulty levels of the items in i category at each step j. Whereas the equation scanned in dichotomous is reduced to (6)

$$P (X_{ni} = x_{ni}) \frac{exp (x_{ni} w_{ij}(\beta_n - \delta_i))}{1+exp (x_{ni} w_{ij}(\beta_n - \delta_i))} \tag{6}$$

The difficulty levels in the QUEST Program are clearly expressed in the threshold value. The threshold value is calculated based on the value of τ which represents the ability levels required by the test (case / person) with the highest chance of 0.50.

3.   RESULTS AND DISCUSSION
3.1. Fitting of Rasch model (Parameter logistic model)

The Logistic parameter model displays the output of suitability of items in detecting learning difficulties through self-regulation strategies. The compatibility of items developed both based on INFIT MNSQ or INFIT T is in accordance with the requirements (Table 1). This is in accordance with the results of the IRT analysis that the distribution of all items has been fit according to the Rasch Model which is in the range of scores from 0.77 to 1.30. This means that the construct of items is appropriate and effective for measuring self-regulation indicators.

The output from the QUEST program provides information about the feasibility of items analyzed according to the Rasch model. Table 1 illustrates the output that informs about the distribution of items based on their compatibility with the Rasch model. The QUEST output provides reference options for the suitability of models developed based on INFIT MNSQ criteria or INFIT T.

Table 1 shows the output distribution profile of items specifically for fit criteria based on MNSQ INFIT. All items that have been used are within the range received (fit by the Rasch model) according to the INFIT MNSQ and INFIT criteria. There are no items that suggest to be improved because they meet the validity criteria. This means that the output produced can be fully used to interpret the probability of responding self-regulation behavior. In addition, it can be stated that the items developed have been feasible to detect learning difficulties through indicators of self-regulation ability. This finding is very important as a

prerequisite for detecting learning difficulties for prospective physics teachers based on self-regulation skills. Analysis of the Rasch model is able to inform about the characteristics of the test items formed into the same metrics.

Table 1. Items that are fit with the Rasch model

Items	Infit MNSQ	Infit t	Outfit t	Criterion
1	1.14	1.2	1.0	*Fit*
2	1.01	0.1	0.2	*Fit*
3	1.10	0.8	0.9	*Fit*
4	1.06	0.6	0.5	*Fit*
5	0.86	-1.2	-0.9	*Fit*
6	1.15	1.2	1.4	*Fit*
7	0.86	-1.2	-1.0	*Fit*
8	0.84	-1.3	-0.9	*Fit*
9	0.86	-1.2	-1.0	*Fit*
10	1.07	0.6	0.5	*Fit*
11	1.01	0.1	0.0	*Fit*
12	1.12	1.0	0.8	*Fit*
13	0.90	-0.8	-0.7	*Fit*
14	1.06	0.6	0.5	*Fit*
15	1.04	0.4	0.4	*Fit*
16	0.94	-0.5	-0.4	*Fit*
17	0.82	-1.5	-1.2	*Fit*
18	0.87	-0.9	-0.1	*Fit*
19	0.92	-0.6	0.1	*Fit*
20	0.84	-1.3	-1.2	*Fit*

Criteria for fit items: 0, 77 ≤ INFIT MNSQ ≤ 1, 33 [23] or -2 ≤ INFIT T≤ 2 [17].

According to Lowenthal [35], test items in quantitative analysis must have the reliability to measure the degrees of consistency. This means that, good measurements must present a high degree of reliability if the scores are consistent [36]. The Rasch model is still an analytical tool that can measure the accuracy of the validity and reliability of research instruments, even testing the suitability of people and items simultaneously. According to Christensen [37], Rasch modeling can produce measurement instruments that are better and more accurate than others.

3.2. Case estimates for identifying self-regulating ability

The quest output displays details of skill levels from each case based on the score estimates. High estimated scores can be interpreted that the self-regulation skills of prospective teachers are good. This means that self-regulation is proportional to an increase in learning motivation and affect learning difficulties. The score estimates in the IRT require respondents who have high ability (Estimates scores> 1.00), medium (-1.00 ≤ Estimates scores ≤1.00), and low (Estimates scores<-1.00) as shown in Table 2.

Table 2. Physics teacher candidates based on their level of self-regulation ability.

No	Case/Person estimates			Levels of self-regulation Skills
	Case estimates (person abilities)	Sum of cases	Persentages (%)	
1	Est > 1.00	42	33	High ability
2	-1.00 ≤ Est ≤1.00	64	51	Medium ability
3	Est < -1.00	20	16	Low ability

Table 2 shows the number of percentage of physics teacher candidates who possess self-regulation skills for the categories of: high, medium, and low abilities. The percentage of the number of prospective teachers who have high self-regulation skills is 33%, and then most of them are in the medium category ie 51%, while 16% are of low ability. It was found that most of the physics teacher candidates had the ability to categorize 'medium'. This means that, the use of alternative learning strategies can improve one's self-regulatory abilities and minimize learning difficulties. According to Bembenutty, et al, [38] that learners who have high self-regulation can overcome learning difficulties because motivation is getting better. This motivation is a psychological cause that influences learning difficulties [39-41].

For prospective teachers, self-regulation is very important to be trained. Good self-regulation can help prospective teachers become accustomed to completing administrative, teaching, and research assignments. A total of 126 cases that have been identified show the results of the estimation that 51% of prospective physics teachers have medium category self-regulation, 33% for high, and 16% with low self-regulation (Table 2). These results have proven that self-project has greatly facilitated the assumption of the Rasch Model [42, 43]. The difference in self-regulation ability was finally identified through the results of the Rasch Model analysis to be interpreted. Joyce [44] said that the implementation of learning strategies is an alternative solution to stimulate learners' self-regulation skills.

4.   CONCLUSION

Learning difficulties are a psychological cause that can disrupt student self-regulation in learning. The Rasch Model is feasible to measure self-regulation skills through a self-sustaining project strategy. The appropriateness between learning strategies and learner characteristics has contributed positively for self-regulation.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Campanario, J. M., Martin, B., Campanario, J. M., & Martin, B. "Challenging dominant physics paradigms," *Against the Tide: A Critical Review by Scientists of How Physics and Astronomy Get Done*, vol. 18(3), pp. 11, 2008.
[2]   Retnawati, H., Arlinwibowo, J., Wulandari, N. F., & Pradani, R. G. "Teachers'difficulties and strategies in physics teaching and learning that applying mathematics," *Journal of Baltic Science Education*, vol. 17(1), 2018.
[3]   Ornek, F., Robinson, W. R., & Haugan, M. P. "What makes physics difficult?," *International Journal of Environmental and Science Education*, vol. 3(1), pp. 30-34, 2008.
[4]   English, M. C., & Kitsantas, A. "Supporting student self-regulated learning in problem-and project-based learning," *Interdisciplinary journal of problem-based learning*, vol. 7(2), pp. 6, 2013.
[5]   Baumeister, R. F., & Vohs, K. D. "Self-Regulation, ego depletion, and motivation," *Social and personality psychology compass*, vol. 1(1), pp. 115-128, 2007.
[6]   Oyserman, D., Lewis Jr, N. A., Yan, V. X., Fisher, O., O'Donnell, S. C., & Horowitz, E. "An identity-based motivation framework for self-regulation," *Psychological Inquiry*, vol. 28(2-3), pp. 139-147, 2017.
[7]   Baghaei, P., & Carstensen, C. H. "Fitting the mixed rasch model to a reading comprehension test: identifying reader types," *Practical Assessment, Research & Evaluation*, vol. 18(5), pp. 1-13, 2013.
[8]   Zimmerman, B. J., & Kitsantas, A. "Acquiring writing revision skill: Shifting from process to outcome self-regulatory goals," *Journal of Educational Psychology,* vol. 91, pp. 241-250, 1999. http:// dx.doi.org/10.1037/0022-0663.91.2.241.
[9]   Zimmerman, B. J. "Attaining self-regulation: A social cognitive perspective," In M. Boekaerts, P. Pintrich, & M. Zeidner, *et al*, *Handbook of self-regulation*, pp. 13-39, 2000.
[10]   Zimmerman, B. J., & Kitsantas, A. "The hidden dimension of personal competence: Selfregulated learning and practice," In A. J. Elliot & C. S. Dweck, *et al*, *Handbook of competenceand* motivation, pp. 204-222, New York, NY: Guilford Press, 2005.
[11]   Zimmerman, B. J., & Schunk, D. H. "Motivation: An essential dimension of self-regulated learning," In D. H. Schunk & B. J. Zimmerma, *et al*, *Motivation and self-regulated learning:Theory, research, and* applications, pp. 1-30, Mahwah, NJ: Lawrence Erlbaum Associates, 2008.
[12]   Zimmerman, B. J. "From cognitive modeling to self-regulation: A social cognitive career path," *Educational Psychologist*, vol. 48(3), pp. 135-147, 2013. http://dx.doi.org/10.1080/00461520.201 3.794676.
[13]   Zimmerman, B. J., & Schunk, D. H. *et al*. "*Self-regulated learning and academic achievement: Theory, research, and practice*," Springer Science & Business Media, 2012.
[14]   Bembenutty, H., Cleary, T., & Kitsantas, A., "*Applications of self-regulated learning appliedacross diverse disciplines: A tribute to Barry J. Zimmerman*," Charlotte, NC: Information Age Publishing, 2013.
[15]   Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., ... & Wittrock, M. C. "A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives, abridged edition," *White Plains, NY: Longman,* 2001.
[16]   Andrich, D. "Probabilistic models for some intelligence and attainment tests," 2005.
[17]   Baghaei, P., & Carstensen, C. H. "Fitting the mixed Rasch Model to a reading comprehension test: Identifying reader types," *Practical Assessment, Research & Evaluation*, vol. 18(5), pp. 1-13, 2013.
[18]   Embretson, S. E. "Mixed Rasch Models for measurement in cognitive psychology," In M. von Davier, & C. H. Carstensen, *et al, Multivariateand mixture distribution Rasch models: extensions andapplications*, pp. 235-253, New York: Springer Verlag, 2007.

[19] Fischer, G. H. "Unidimensional linear logistic Rasch Models," In *Handbook of modern item response theory*, pp. 225-243, Springer, New York, NY, 1997.

[20] Andersen E.B. "The rating scale model," In: van der Linden W.J., Hambleton R. K, *et al*, *Handbook of Modern Item Response Theory*. Springer, New York, NY, 1997.

[21] Andersen, E. B. "A goodness of fit test for the Rasch Model," Psychometrika, vol. 38(1), pp. 123-140, 1973.

[22] Dehqan, A., Yadegari, F., Asgari, A., Scherer, R. C., & Dabirmoghadam, P. "Development and validation of an Iranian Voice Quality of Life Profile (IVQLP) based on a classic and Rasch Rating Scale Model (RSM)," *Journal of Voice*, vol. 31(1), pp. 113-119, 2017.

[23] Baghaei, P., Yanagida, T., & Heene, M. "Development of a descriptive fit statistic for the Rasch model," 2017.

[24] Henninger, M. "A novel varying threshold IRT approach to accounting for response styles," *Psychometric modeling as a tool to investigate heterogeneous response scale use*, pp. 129, 2019.

[25] Andrich, D. "Probabilistic models for some intelligence and attainment tests," 2005.

[26] Dehqan, A., Yadegari, F., Asgari, A., Scherer, R. C., & Dabirmoghadam, P. "Development and validation of an Iranian Voice Quality of Life Profile (IVQLP) based on a classic and Rasch Rating Scale Model (RSM)," *Journal of Voice*, vol. 31(1), pp. 113-119, 2017.

[27] Baghaei, P., Yanagida, T., & Heene, M. "Development of a descriptive fit statistic for the Rasch model," 2017.

[28] Henninger, M. "A novel varying threshold IRT approach to accounting for response styles," *Psychometric modeling as a tool to investigate heterogeneous response scale use*, pp. 129, 2019.

[29] Bond, T. G., & Fox, C. M. "*Applying the Rasch Model: Fundamental measurement in the human sciences*," Psychology Press, 2013.

[30] Dogan, E. "An Application of the partial credit IRT model in identifying benchmarks for polytomous rating scale instruments," *Practical Assessment, Research & Evaluation*, vol. 23(7), 2018.

[31] Bonanomi, A., Olivari, M. G., Mascheroni, E., Gatti, E., & Confalonieri, E. "Using a multidimensional rash analysis to evaluate the psychometric properties of the motivated strategies for learning questionnaire (mslq) among high school students," *TPM: Testing, Psychometrics, Methodology in Applied Psychology*, vol. 25(1), 2018.

[32] Jong, C., Hodges, T. E., Royal, K. D., & Welder, R. M. "Instruments to measure elementary preservice teachers' conceptions: an application of the rasch rating scale model," *Educational Research Quarterly*, vol. 39(1), pp. 21, 2015.

[33] Anthony, C. J., DiPerna, J. C., & Lei, P. W. "Maximizing measurement efficiency of behavior rating scales using item response theory: An example with the social skills improvement system—teacher rating scale," *Journal of school psychology*, vol. 55, pp. 57-69, 2016.

[34] Linacre, J. "Sample size and item calibration stability," *Rasch Mes Trans.*, vol. 7, pp. 328.

[35] Loewenthal, K., & Lewis, C. A. "*An introduction to psychological tests and scales*," Psychology press, 2018.

[36] Nielsen, T. "The intrinsic and extrinsic motivation subscales of the Motivated Strategies for Learning Questionnaire: A Rasch-based construct validity study," *Cogent Education*, vol. 5(1), pp. 1504485, 2018.

[37] Christensen, K. B., & Kreiner, S. "Item fit statistics," *Rasch Models in health*, pp. 83-104, 2012.

[38] Bembenutty, H., Cleary, T., & Kitsantas, A., "*Applications of self-regulated learning applied across diverse disciplines: A tribute to Barry J. Zimmerman*," Charlotte, NC: Information Age Publishing, 2013.

[39] Zimmerman, B. J., & Schunk, D. H. "Motivation: An essential dimension of self-regulated learning," In D. H. Schunk & B. J. Zimmerman, *et al*, *Motivation and self-regulated learning: Theory, research, and applications*, pp. 1-30, Mahwah, NJ: Lawrence Erlbaum Associates, 2008.

[40] Fisher, O., O'Donnell, S. C., & Oyserman, D. "Social class and identity-based motivation," *Current opinion in psychology*, vol. 18, pp. 61-66, 2017.

[41] Oyserman, D. "The essentialized self: implications for motivation and self-regulation," *Journal of Consumer Psychology*, vol. 29(2), pp. 336-343, 2019.

[42] Adams, R. J., & Kho, S. T. "Acer quest version 2.1," *Camberwell, Victoria: The Australian Council for Educational Research*, 1996.

[43] Wright, B. D., & Masters, G. N. "*Rating scale analysis*," MESA press, 1982.

[44] Joyce, B. W., & Weil, M. M. dan Calhoun, E. "Models of teaching," 2009.

BIOGRAPHIES OF AUTHORS

Habibi is Lecturer of physics education at IKIP Mataram UNDIKMA, Indonesia. He is a PhD student at Yogyakarta State University. Research and publication are interested in the topics of developing learning and instruction, self-regulated learning, project based learning.

Jumadi is a Professor in the mathematics and science education faculty. Research and publication on topics of instructional technology, instructional, evaluation, learning models development. He has variously presented at seminars and published in national and international journals

Mundilarto is a Professor of Physics Learning and Evaluation, Yogyakarta States University. Research and publication on topics of learning instructional, evaluation, learning development. He has variously presented at seminars and published in national and international journals