PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 77, No. 6, 2019
790

THE ESTIMATION OF TEACHING PROFESSION EDUCATION (TPE) PARTICIPANTS' ABILITY IN INDONESIA: AN ASSESSMENT MODEL WITH PARTIAL SCORING

Lian Gafar Otaya, Badrun Kartowagiran, Heri Retnawati
Yogyakarta State University, IAIN Sultan Amai Gorontalo, Indonesia
E-mail: lian055pasca2016@student.uny.ac.id, lianotaya82@iaingorontalo.ac.id,
kartowagiran@uny.ac.id, heri_retnawati@uny.ac.id

Abstract

The ability of TPE participants in arranging lesson plans is essential to be mastered as an indicator of pedagogical competence which should be possessed before conducting a teaching process in the class. This research aimed to analyze the estimation of TPE participants' ability in composing a lesson plan which uses the assessment with partial scoring. This research used quantitative descriptive explorative approach. Data were collected through document study of lesson plan made by TPE participants assessed by lecturers in a workshop. In the sample were 236 respondents selected by using purposive sampling technique. The data analysis was performed to estimate participants' skill using Item Response Theory through Partial Credit Model (PCM). The estimation result produced the interpretation of items with 5 categories of scoring 0, 1, 2, 3, 4 and has 4 thresholds such as the threshold of category 0 to category 1, category 1 to category 2, category 2 to category 3, and from category 3 to category 4. Besides that the information function value of the instrument and Standard Error of Measurement based on the analysis obtained items in the instrument which have higher information function value compared to the mistakes of the measurement or the estimation of measurement which has relatively small errors, so it can be stated that learning model is using accurate partial scoring to estimate the ability of TPE participants in composing lesson plans.

Keywords: *teaching profession, partial scoring, item response model, partial credit model.*

Introduction

The quality of teachers becomes global attention in every developed country and is strongly related with the critics and concensus which requires the standard quality of teachers and how to improve it because it directly has positive effect towards students' learning result and the quality of education of a country (Gareis & Grant, 2014; Good, 2008; Goodwin, 2010; Retnawati, Djidu, Apino, & Anazifa, 2018; Stronge, 2018; Trygstad, Banilower, Smith, & Nelson, 2014). However, the condition in Indonesia shows disparity between quantity and quality of teachers with the quality of education. The development and recent condition in Indonesia is in emergency situation. As stated by Baswedan (2014), based on the mapping result by the Minister of Education and Culture towards 40.000 schools and in 2012 there were 75% of schools in Indonesia which did not meet the standard of education service. The result of teacher competency test in 2012 toward 460.000 teachers only achieved an average value of 44.5 from the standard expected which is 70. The mapping of the access and the quality of education in 2013 and 2014 conducted by *The Learning Curve-Pearson* shows that Indonesia is

https://doi.org/10.33225/pec/19.77.790 ISSN 1822-7864 (Print) ISSN 2538-7111 (Online)

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 77, No. 6, 2019

included to the countries with low performances. The mapping of the quality of high education by 21 universities in 2013 showed that the position of Indonesia is 49 of 50 countries. Besides that, physical abuse in educational environment always becomes the news, moreover the sexual abuse often occurs in school environment. Bhakti and Maryani (2017) revealed data of teacher competency test in 2015 towards 2.430.427 of teachers in Indonesia which nationally has not reached the average target 53,05 of 55 obtained from average professional ability 54,77 and the pedagogical competency average value is 49.84. There are 7 provinces which reach national average scores, those provinces are Special Region of Yogyakarta (62.48), Central Java (59.10), DKI Jakarta (58.44), East Java (56.73), Bali (56.13), Bangka Belitung (55.13), and West Java (55.05) which averagely only reach the score 41.96 for the province outside Java.

The other problems faced by Indonesia are the disparity between quantity and quality of teachers previewed by the qualification of education and educator certificate, one of them as stated by Fahdini, Mulyadi, Suhandani, and Julia (2014) where certified teachers should improve the self-quality and there is no difference before and after certification. Mangkunegara and Puspitasari (2015) stated that there is no appropriate academic qualification of teachers with the subject in the class. Krisna (2017) stated that there is the inequality of the distribution of teachers in the regions causing many schools lack of qualified teachers especially in the rural areas. Based on the survey of *Program for international student assessment* (PISA) in 2015 announced in 6 December 2016, Indonesia is in the position of 64 from 72 surveyed countries. Therefore, the survey conducted by *Progress in International Reading Literacy Study* (PIRLS) and *Trend in International Mathematics and Science Study* (TIMMS) also do not significantly improve the position of Indonesia. Besides that, every moral deviation among students, teachers who are considered to be less professional in teaching (Antoro, 2017). Data shows the problems in education become very fundamental and urgent to be seriously handled, because it has positive and significant correlation with the problems faced by the teachers in Indonesia.

Many policies had been taken by government in improving qualification of teachers and stabilizing the competency of pre-service teachers in Indonesia. The assessment and the development of teachers' carrier in Indonesia are also included as the essential components in the continual development of profession (Chang et al., 2013). The government of Indonesia has issued various policies about the constitution in order to improve the competency of teachers and pre-service teachers in Indonesia through the teaching profession education (TPE). The program is urgently needed in preparing professional teachers and also is a strategy in education system which can improve teachers' performance. This program is also the only way to develop teacher profession, so their performance can be better in improving students' achievement (Hotimah & Suyanto, 2017; Mizell, 2010).

The implementation of teaching profession education (TPE) in Indonesia implements a consecutive model in accordance with Constitution Number 20 Year 2003 regarding national education system and the regulation of government number 74 year 2008 about teacher and it is reinforced in Constitution Number 12 Year 2012 about high level education. TPE is conducted referring to the Regulation of Ministry of National Education Number 8 Year 2009 about education program of pre-service teacher where recently that regulation has been changed into the Regulation of Ministry of Education and Culture Number 87 Year 2013 as a strong legal protection in improving the quality of teachers in Indonesia. In the Regulation of Minister of Education and Culture Number 87 Year 2013 article 1 (2), it is mentioned that "Teacher Profession Education Program which is further called as TPE program is an educational program conducted to prepare the graduates of bachelor degree in education program and non-education program who have potency to be good teachers in all levels of education".

The commitment of government is shown by the reissuing of the policy through the Regulation of Minister of Research, Technology, and High-Level Education Number 55 Year 2017 regarding education standard of the teacher. In that regulation, Chapter I article 1 (5) it

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 77, No. 6, 2019

is mentioned that: "Teaching Profession Education Program which is furthermore called as TPE. TPE is an education program conducted after bachelor's degree or applied by a bachelor to obtain educator certificate in all levels of education". It emphasizes that to be a teacher, a graduate of a bachelor program should join TPE program, and after passing the program he/she will be given educator certificate and will become a teacher. It means that, teacher is a profession which requires special requirements.

TPE in Indonesia is also a response of government towards the national needs of teacher which focuses on the deep mastery towards the field taught, knowledge and pedagogical skill, and also has social skill and good personality, so that in TPE develops four basic competencies which should be possessed and mastered by professional teachers, those competencies are: pedagogic competency through learning activity, social competency and personality through dormitory activity and professional competency through field teaching practice (Hotimah & Suyanto, 2017; Wahyudin, 2016). Therefore, to be a professional teacher, it is essential to join TPE program and put teacher training institution in Indonesia as an institution which is most responsible in the realization.

TPE program is conducted by teaching training universities which meet the requirements proposed by the Minister of Research, Technology, and High Education in Indonesia. The requirements which should be fulfilled by teaching training universities are the following: the accreditation of study program at least should be B, the commitment of teaching training universities, the quality of supporting facilities, dormitory facility, having the improvement and the development of instructional activities, partnership program with the partner school with the accreditation at least B and meet the qualification to conduct field practice. It shows that teaching training universities which conduct TPE in Indonesia should meet the equal quality standard of education in realizing the competence achievements of TPE participants which describe the knowledge, skill, and behavior in conducting the professional duty as a teacher with some qualifications required in the form of teachers' competences. Those four teacher competences are: pedagogical competence, personal competence, social competence, and professional competence. From those competences, one of them is strongly related to teacher profession. The main aspect mostly concerned is the learning process and pedagogical competence especially in composing lesson plans. The success of learning process depends on the lesson plans as the guidance, so the ability of TPE participants in composing lesson plans should be estimated.

Estimating the ability of TPE participants in composing lesson plans is essential to be conducted starting from formulating competence indicator and learning output, organizing learning materials, process, source, media, assessment, and learning evaluation and implementing *Techno-Pedagogical Content Knowledge (TPACK)* principle, as an effort to improve the quality of learning process and the TPE participants' mastery of pedagogical competences as a teacher. In estimating that ability, the high accuracy is totally needed to avoid the risk or the error in the measurement. The aspects that should be noticed are such as the components measured in the lesson plan and also the scoring model used to give accurate and credible results. According to the theory of the classic scoring measurement, it is usually used based on some steps and the score per item of participants is obtained by summing the score in each step (Istiyono, Mardapi, & Suparno, 2014a; Retnawati, 2011). This assumption is developed in the theory of item response which states that an item should follow the pattern of partial credit and it is expected to have higher ability (Istiyono, Mardapi, & Suparno, 2014b; Widhiarso, 2010). Based on that assumption, it estimates participants' ability in composing lesson plans which can use partial scoring.

Partial scoring is a model which is simple and easy to use. Partial scoring shows the steps to correctly accomplish steps of the component and each step that is done correctly is given the score 1 and that is done incorrectly is given the score 0, for example: true-false, yes-no, do-not

https://doi.org/10.33225/pec/19.77.790 ISSN 1822-7864 (Print) ISSN 2538-7111 (Online)

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 77, No. 6, 2019

do an activity, there is-none and others (Retnawati, 2016; Safaruddin, Anisa, & Saleh, 2018). Based on the characteristics of the partial scoring, the question that arises as the main problem in this study is whether this scoring model can provide the estimated results of item parameter and the ability of participants to prepare learning plans according to the components that must be assessed, so that the accuracy of the scoring model can be known in estimating the ability of TPE participants to follow up on improvements.

Research Problem

Some studies have been conducted related to the implementation of TPE in Indonesia, one of them like what was conducted by Hidayah (2013) at Semarang State University who found that the implementation of this program has given good output for the participants. Therefore, a research conducted by Basri (2015) revealed that this program totally needs to be previewed and really reach the goals targeted by participants. From a research conducted by Indriyani and Ismandari (2015) at Yogyakarta State University that involved 255 students, it is found that most of the students (50.20%) hesitate toward their participation in this program, therefore by the evaluation or the assessment of this program there are 48.23% of students who hesitate. A research conducted by Ningsih, Fatchan, and Susilo (2016) in the implementation of program of TPE SM3T of geography teachers in Semarang State University shows that the difference of learning device discussed in the workshop and there are some participants who should have remedial because have low scores. This research was also reinforced by Setiajid, Susanti, and Ngabiyanto (2018) who state that the perfection of the management of quality of service should be done to measure the achievement.

The research previously conducted showed that TPE program is an activity which really depends to various internal component systems. It is reinforced by Zamakhsari (2009) who stated that the achievement of TPE program is strongly determined by many components such as: student admission system, curriculum, learning process, assessment, and the quality guarantee of that program. From those aspects, there is one aspect which is essential and urgent in affecting TPE program in every teacher training university in Indonesia. That is the assessment model which can reveal the mastery of students' competence and after joining the program, although there is no specific research about the development of assessment model in TPE program in Indonesia including in estimating students' ability to arrange learning plan as one output of pedagogic competency which should be possessed by participants before the learning process in class. Therefore, this research is very important, urgent, and interesting to be learned, because one of the characteristics of this scoring based on the approach of item response theory polytomus model is the higher score which shows the bigger ability than the lower category score because the score obtained ranges from only 0 – 1, which possibly causes someone to gain the lowest and the highest score. In addition, every stage has different difficulty level and every stage is scored 1, because it has same difficulty level. Otherwise, the stage which cannot be conducted is scored 0. Therefore, the scores obtained are collected based on the response in every stage, so it is possible to describe the stages mastered by TPE participants. The stages measured from TPE participants are usually correct in a certain stage, but they are incorrect in the next stage or others. Therefore, partial scoring is an option which has opportunity to provide information about the ability of TPE participants. The higher the parameter of TPE participants' ability, the bigger the chance of the participant in correctly doing each stage in the items assessed in the lesson plan.

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 77, No. 6, 2019

Research Focus

The research was to describe the estimation of TPE participants' ability in composing a lesson plan which uses the assessment with partial scoring. The estimation was conducted using the approach of item response theory to find the position of TPE participants in composing lesson plan based on the ability scale, so the ability of participants was known and can be compared among them especially in the scoring based on theta scale.

Research Methodology

Design

The research was an explorative descriptive research which was the result of Research and Development (R & D). This research described the result of TPE participants' ability in composing lesson plan in a workshop, field teaching practice and performance test based on the data of the assessment from the supervisor and the examiners. Data were collected from December 2018 until February 2019 and the analysis was conducted from March until April 2019.

Population and Sample

The population in this research was all subsidized Pre-service TPE and TPE Participants in the 2018/2019 academic year held by Yogyakarta State University, Gorontalo State University, and Alauddin Islamic State University. The sample of this research was 236 participants selected using a purposive sampling technique based on some considerations of the representatives as the participants. The sample was selected which corresponded to the number of participants in each study program in the implementation of Subsidized Pre-Service TPE and TPE in Yogyakarta State University, Gorontalo State University, and Alauddin Islamic State University in academic year 2018/2019. In this research, there were 48 Subsidized Pre-Service TPE participants at Yogyakarta State University in some study programs such as Primary Education Study Program, Mathematics Education, English Education, and Automotive Engineering Education involving 21 field supervisors as assessors in 13 schools that are locations of the field teaching practice namely: SDN 1 Jarakan, SDN Giwangan Yogyakarta, SDN Bantul Timur, SDN Percobaan 2, SMPN 1 Yogyakarta, SMPN 6 Yogyakarta, SMAN 1 Yogyakarta, SMAN 2 Yogyakarta, SMAN 6 Yogyakarta, SMKN 2 Depok, SMKN 2 Yogyakarta and SMKN 3 Yogyakarta.

Furthermore, the 85 subsidized pre-service TPE participants at Gorontalo State University in English Education Study Program, and TPE participants in Primary Education Study Program, English Education, and Physical Education, Health and Sports whose assessments were conducted when TPE participants carried out performance test Practices involving 10 examiners as assessors at the school under the supervision of Gorontalo State University namely SDN 84 Kota Tengah, SDN 85 Kota Tengah, Laboratory School of Gorontalo State University, SDN 30 Kota Selatan, SDN 29 Kota Selatan, SMPN 1 Gorontalo, SMPN 2 Gorontalo and SMAN 3 Gorontalo. For TPE participants at Alauddin Islamic State University, there were 103 TPE participants with study programs / subjects including: *Fikih*, Qur'an Hadith, Morals, Arabic, Islamic Cultural History and Teachers of Madrasah Ibtidaiyah whose assessments were carried out when TPE participants conducted workshop and the field teaching practice involving 15 workshop lecturers / field advisers on field teaching practice in MIN 1 Makassar, MI Madani Alauddin Pao-Pao, MI Ash-Shalihin, MTs Guppi Samata, MTs Madani Alauddin Pao-Pao, MA Guppi Samata, MI Ash-Shalihin, MAS Nadhlatul Wathan.

https://doi.org/10.33225/pec/19.77.790 ISSN 1822-7864 (Print) ISSN 2538-7111 (Online)

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 77, No. 6, 2019

Instrument and Procedures

Data were collected using the documentation technique through analysis of the document of lesson plan that TPE participants had compiled during the workshop, field teaching practice and performance test. The instrument used in this research assessed 4 indicators that were described into 25 items that were assessed in detail as follows: the formulation of indicators of achievement of competencies and learning outcomes consisted of 6 items, organizing materials, methods, media and learning resources also consisted of 6 items, organizing processes, learning evaluation and evaluation assessed 6 items, and the implementation of the principles of techno pedagogical content knowledge which assessed 7 items. The distribution of items based on the assessed indicators is presented in Figure 1:

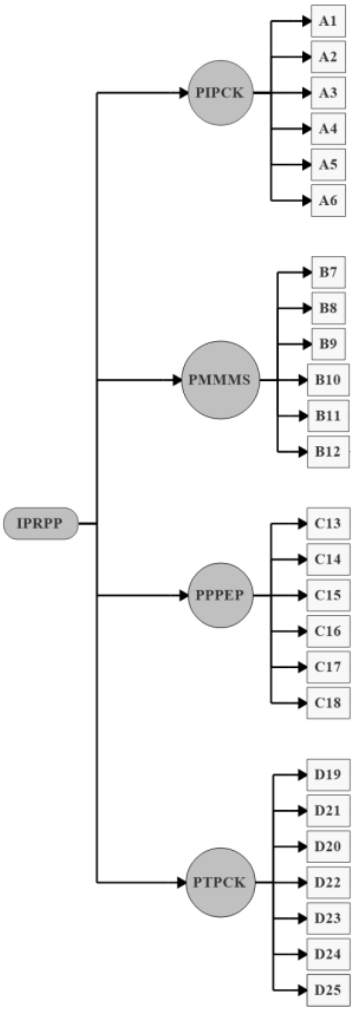

Figure 1. The distribution of item indicator of lesson plan assessment.

Explanation :
IRPP : Lesson Plan
PIPKC : The Formulation of the indicator of competency achievement and learning achievement
PMMMS : The organizing of material, method, media and learning source
PPPEP : The organizing of process, assessment, and learning evaluation
PTPCK : The implementation of *techno pedagogical content knowledge* principle

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 77, No. 6, 2019

The instruments used in this research were accurate and reliable. It was proved by the content validity of seven expert judgment results from the analysis using the Aiken's V table with a *p*-value of .05 for the five categories obtained by the Aiken index V = .75, meaning that each item was valid if it meets the required Aiken index ie > .75 (Aiken, 1985). The results of the analysis in general show that all items were valid as presented in Table 1:

Table 1. The categorization of content validity instrument of 25 items.

Indicator	Table Aiken *V*	
	Valid (> 0.75)	Invalid (< .75)
PIPCK	A1, A2, A3, A4, A5, A6	-
PMMMS	B7, B8, B9, B10,B11, B12	-
PPPEP	C13,C14, C15, C16, C17, C18	-
PTPCK	D19, D20, D21, D22, D23, D24 D25	-
Total / %	25/100%	0 / 0%

Furthermore, to see the consistency or reliability assessment of seven expert judgments, reliability estimation using inter-rater reliability techniques obtained reliability coefficient of .84, the instrument was reliable if the coefficient exceeds the criteria ≥ .70 (Gronlund & Linn, 1965; Nunnally, 1994). It generally shows that all items in the instrument are trusted and can be used in the data collection process.

Data Analysis

The data analysis in this research used the approach of item response theory polytomus model using partial credit model (PCM) with the help of the R Program Software with syntax extended Rasch Modeling (eRM). As stated by Retnawati (2014) The category score on PCM shows the number of steps to complete the item correctly. Higher category scores indicate greater ability than lower category scores. In PCM, if an item has two categories, then equation 2 becomes the Rasch model equation. As a result of this, PCM can be applied to polytomous and dichotomous items. The equation above, can be described based on the number of categories in the item, so that the Rasch model equation is used to analyze the response of items with 4 choices, then obtained categories (j) of 4 (j = 0, 1, 2, 3) equations of probability each individual in the category response function (CRF) is realized in the following equation.

https://doi.org/10.33225/pec/19.77.790 ISSN 1822-7864 (Print) ISSN 2538-7111 (Online)

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 77, No. 6, 2019

Category 0 :

$$P_{10}(\theta) = 0 - P_{i1}^{*}(\theta) = 0 - \frac{\exp[Da_i(\theta - b_{11})]}{1 + \exp[Da_i(\theta - b_{11})]}$$

Category 1 :

$$P_{11}(\theta) = 0 - P_{i1}^{*}(\theta) - P_{i2}^{*}(\theta) = \frac{\exp[Da_i(\theta - b_{11})]}{1 + \exp[Da_i(\theta - b_{11})]} - \frac{\exp[Da_i(\theta - b_{12})]}{1 + \exp[Da_i(\theta - b_{12})]}$$

Category 2 :

$$P_{12}(\theta) = 0 - P_{i2}^{*}(\theta) - P_{i2}^{*}(\theta) = \frac{\exp[Da_i(\theta - b_{12})]}{1 + \exp[Da_i(\theta - b_{12})]} - \frac{\exp[Da_i(\theta - b_{13})]}{1 + \exp[Da_i(\theta - b_{13})]}$$

Category 3:

$$P_{13}(\theta) = 0 - P_{i3}^{*}(\theta) - P_{i4}^{*}(\theta) = \frac{\exp[Da_i(\theta - b_{13})]}{1 + \exp[Da_i(\theta - b_{13})]} - \frac{\exp[Da_i(\theta - b_{14})]}{1 + \exp[Da_i(\theta - b_{14})]}$$

Furthermore, to determine the accuracy of this scoring in estimating the ability of TPE participants to use the value of the information function (item information function) and standard error of estimation (Standard Error Measurement, SEM). According to Retnawati (2014) in item response theory, the value of information function is a method to explain the strength of an item in a test set, the selection of test items, and comparison of several test kits that can express the strength or contribution of studied items in revealing latent trait measured with these instruments and can be known which items are compatible with the model. That the value of information functions has an inverse relationship with the Standard Error of Measurement (SEM), the greater the value of information, the SEM will be smaller or vice versa (Hambleton, Swaminathan, & Rogers, 1991; Retnawati, 2014). If the value of the information function is known, then directly SEM can be obtained using the formula.

$$\text{SEM } \hat{(\theta)} = \frac{1}{\sqrt{I(\theta)}}$$

therefore,
$I(\theta)$: Information function

Research Results

Estimating the ability of TPE participants in preparing learning plans by assessment using a partial scoring model requires a number of stages that must be met according to the component being assessed. At each of these stages the participants' mastery is assessed in meeting the criteria for each stage of the preparation of the learning plan. A partial scoring in the assessment application on the observed aspect is given a score for each step taken even if it does not do all of the steps or may not be sequential but carried out. Partial scoring indicates the number of steps to complete the steps of the component being assessed and this is a characteristic of the partial scoring model. Assessment with partial scoring in this study was conducted by lecturers on TPE participants during workshops / workshops and the implementation of field experience practices in schools by filling in the learning plan assessment instrument sheet and scoring the indicator points / aspects assessed by circling the numbers on score column (score 0, 1, 2, 3, 4 or score 4). Each score has criteria according to the assessment rubric prepared, namely score 0 = no visible descriptors, score 1 = one descriptor visible, score 2 = two descriptors visible, score 3 = three descriptors appear, and score 4 = four descriptors appear.

The response items of the participants in this study were analyzed using the PCM model which assumes that the item difficulty level parameter is the only item characteristic that influences the participants' ability to prepare lesson plans. Based on the results of the PCM model analysis with the help of the R Syntax Extended Rasch Modeling (eRm) Package

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 77, No. 6, 2019

Program to the responses of 236 TPE Program participants from the results of an assessment of learning plans comprising 25 items, resulting in an estimated difficulty level of items or a threshold of more than one. Items with 5 scoring categories in this study: 0, 1, 2, 3, and 4 have 4 thresholds, namely threshold from category 0 to category 1, from category 1 to category 2, from category 2 to category 3, and from category 3 to category 4, because in general items with n categories have (n-1) thresholds. The following results indicate the level of difficulty for items as shown in Table 2:

Table 2. The prediction of threshold using PCM model.

Item	Location	Threshold δ_1	Threshold δ_2	Threshold δ_3	Threshold δ_4
A1	0.36	-0.79	-0.24	0.85	1.66
A2	0.51	-0.60	-0.37	0.92	2.12
A3	0.15	-1.82	-0.05	0.59	1.92
A4	0.26	-1.41	-0.06	1.02	1.53
A5	0.38	-1.54	0.04	0.70	2,32
A6	0.36	-1.35	0.12	0.56	2.12
B7	0.37	-1.12	0.28	0.29	2.04
B8	0.37	-1.17	-0.18	1.00	1.83
B9	0.36	-1.30	0.20	0.54	2.00
B10	0.47	-0.94	-0.08	0.71	2.22
B11	0.36	-0.67	-0.42	0.76	1.79
B12	0.45	-1.09	0.14	0.97	1.78
C13	0.58	-0.93	0.06	1.08	2.13
C14	0.50	-1.33	0.16	0.83	2.34
C15	0.53	-0.96	0.01	0.94	2.12
C16	0.44	-1.64	0.23	0.78	2.39
C17	0.47	-1.24	0.05	1.28	1.81
C18	0.43	-0.97	0.31	0.61	1.79
D19	0.71	-0.36	0.51	0.91	1.78
D20	0.51	-1.01	-0.05	1.09	2.03
D21	0.76	-0.19	0.29	1.08	1.86
D22	0.55	-0.86	0.06	1.09	1.92
D23	0.81	-0.03	0.31	0.83	2.12
D24	0.69	-0.29	0.15	0.84	2.07
D25	0.56	-0.73	0.41	0.82	1.75

Table 2 obtained information that the location parameters of each item varied from 0.15 to 0.81. In addition, in the threshold parameter sebanyaki there are four groups or four intersections. This is a parameter for the level of difficulty participants get a certain score when responding to item i. Judging from the chance of achieving the score, the threshold parameter coefficient δ_i for each category is different. The higher the achievement category, the higher the threshold coefficient δ_i. When viewed at the level of difficulty, the items obtained varying levels of difficulty found in each item and in each category. The findings showed that there are four categories or four intersections which are the level of difficulty. The research findings showed

https://doi.org/10.33225/pec/19.77.790 ISSN 1822-7864 (Print) ISSN 2538-7111 (Online)

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 77, No. 6, 2019

that the easiest category is on the ability scale of -2.06 (category 1 on item 2) while the most difficult category is on the ability scale of 3.0 (category 4 on item 15). This shows that the level of difficulty in the assessment instrument of the learning plan implementation moves in the range of abilities. Another thing that can be explained based on the results of item analysis with the Partial Credit Model is the Item characteristic curve (ICC) which explains the relationship of the ability of TPE participants to an underlying latent item and tool. ICC is described to make it easier to understand the relationship between each threshold $^{\delta}$i which is the level of difficulty with the ability of participants to reach a certain score or category. Following is the description of two examples of item characteristic curve namely item 1 (A1) and item 10 (B10) comprising 25 item instruments of assessment of lesson plan in the partial scoring in the Figure 2:

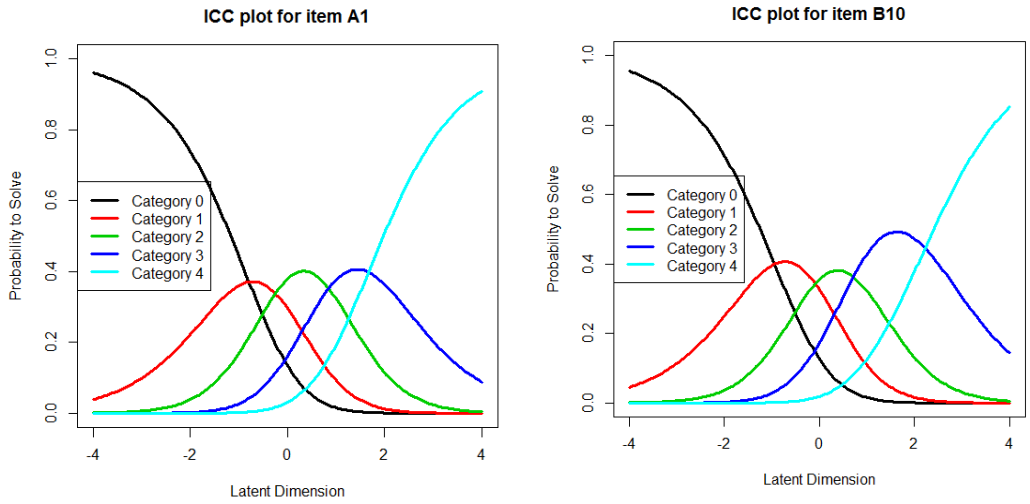

Figure 2. Item characteristic curve, ICC Item A1 and Item B10.

Figure 2 if related to the results of item calibration in Table 1, it can be explained that basically item 1 (A1) has a location parameter of 0.47 with a *threshold* $^{\delta}_{1}$ of -0.79, *threshold* $^{\delta}_{2}$ of -0.24, *threshold* $^{\delta}_{3}$ of 0.85, and *threshold* $^{\delta}_{4}$ which is 1.66. Graphically the threshold $^{\delta}$i can be interpreted as the intersection of the curves of each category. From this figure, it can be explained that to reach category 2 or to get a score of 2 on item 1 (A1), the required ability ()is around -0.24 to 0.85. Furthermore, in item 10 (B10) if related to the results of the caliber items in table 1, it can be explained that basically this item 10 has a location parameter of 0.47 with a threshold parameter *threshold* $^{\delta}_{1}$ of -0.94, *threshold* $^{\delta}_{2}$ of -0,08, *threshold* $^{\delta}_{3}$ of 0,71, and also *threshold* $^{\delta}_{4}$ which is 2.22. Graphically the threshold $^{\delta}_{i}$ can be interpreted as the intersection of the curves of each category. From this figure, it can be explained that in order to reach category 2 or to get a score of 2 on item 10 (B10), the required ability () is around -0.08 to 0.71. In addition to the item characteristic curve, the other thing that can be explained from the quality of the partial scoring lesson assessment instrument for the lecturer evaluator is the value of the instrument's information function. The information function basically shows the extent to which the instrument that has been developed can provide maximum information if it is imposed on certain capabilities. The following is the value of the information function (NIF) assessment instrument for implementing a partial scoring of lesson plan (RPP) for lecturer assessors that is linked to the Standard Error of Measurement (SEM) in Figure 3:

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 77, No. 6, 2019

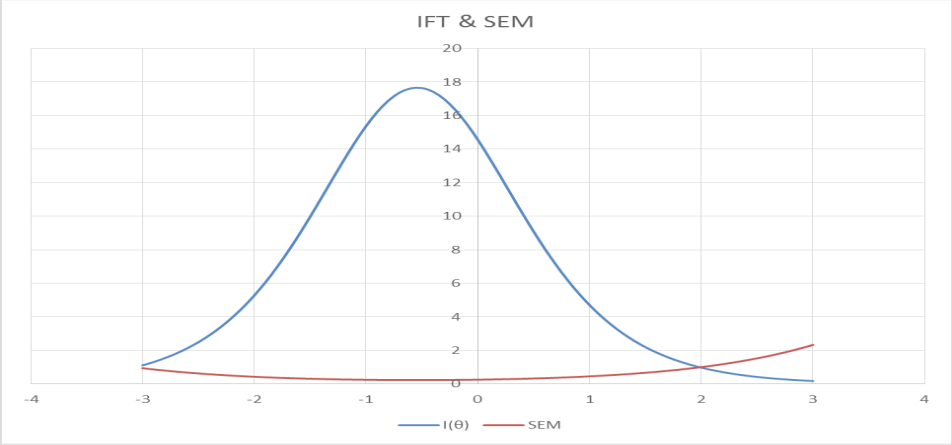

Figure 3. The relation between information function and standard error of measurement.

Figure 3 presents the information function relationship and SEM from the accumulation of 25 instrument items assessing the ability of TPE participants to prepare a lesson plan wherein the curve for the information function is indicated by a curved line going up and the opposite SEM. The two graphics of this function converge on a capability scale of -3 and 1.9. Between those two abilities, the instrument has a higher information value compared to the measurement error. Conversely, when the capability scale is less than -3 and more than 1.9, then this instrument has a measurement error that is greater than the information it provides. Another thing that can be explained is the value of the instrument information function of 17.8 on the ability scale () -0.6. The value of the information function has an inverse relationship with SEM. Therefore, the greater the information value, the smaller SEM will be or vice versa. Based on this opinion, it can be explained that by knowing the value of the information function of 17.8, the SEM coefficient is obtained at 0.23. This means that the items in the instrument have a higher information function value compared to the measurement error.

Discussion

Many new measurement theories have developed over the past decade that can provide detailed information about testing latent attributes or abilities that a person has and are relevant for assessing complex assessments and for evaluating the accuracy of the resulting score (Templin & Hoffman, 2013). In the measurement theory that developed at this time there are two methods of estimation namely the classical test theory and item response theory (modern theory) or more popularly known as the Item Response Theory (IRT). This method assumes that latent variables are represented by a *unidimensional* continuum and can provide precise and detailed information about latent attributes or abilities that a person has and can equalize scores on tests (de Ayala, 2009; Hu, Qin, Sullivan, & Templin, 2017; Sinharay & Haberman, 2014).

Scoring in the IRT model, there are two models, namely the dichotomy and the political model. In the *politomus* model, each scoring model is used to determine the level or category achieved by participants based on the responses given and can be categorized into nominal and ordinal item response models depending on the characteristic assumptions about the data (Retnawati, 2014). The model is very dependent on the characteristics of the parameters and the suitability of the model used, because all response scoring models on the *politomus* are based

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 77, No. 6, 2019

on the assumption that the response to an item depends on the ability of test takers measured using a *unidimensional* scale (Retnawati, 2014; Retnawati & Munadi, 2013). The *politomus* scoring models often used are partial credit model (PCM), generalized partial credit model (GPCM), and graded response model (GRM). Therefore, to find the accuracy of the assessment model in estimating the ability of TPE participants in preparing a learning implementation plan in this study, it was analyzed by PCM using two parameters namely the value of the information function and the estimated standard error. Given the PCM model is a scoring model that evaluates responses on a step-by-step basis. So that, in one PCM scoring process, a number of categories are obtained. Besides that, the threshold for achieving a higher category in scoring this model is not always greater than the previous threshold.

The findings were obtained that the ability of TPE participants in composing a learning implementation plan that was assessed using partial scoring with a rating scale of 0, 1, 2, 3, and 4. The finding of this research was the empirical assessment using a partial scoring of the ability of TPE participants in preparing learning plans which showed that the ability of TPE participants was dominated by the ability () from 1 to 3. The finding was related to the finding of the research conducted by Retnawati and Munadi (2013) which stated the ideal ability parameter was at least 1 or more. It shows that the ability of TPE participants in preparing learning plans are showing good results. In addition, this assessment model had adequate information value and relatively small measurement errors, so that it is accurate in estimating the ability of TPE participants. This finding is in line with the opinion of Retnawati (2014) that the information function has a relationship with latent trait or ability to be measured. This opinion is related to the empirical finding of this study, so it can be explained that the assessment model with partial scoring shows the point of the maximum information function at a higher ability (positive theta). It also means that a partial scoring model is more suitable for TPE participants with high ability, because it will provide maximum information if it is imposed on highly capable participants. This is also supported by measurement error (Standard Error of Measurement) which shows a relatively small value or inversely proportional to the value of the information function, so that it shows the measurement error in the assessment model with an accurate partial scoring in estimating the ability of TPE participants to prepare a learning implementation plan.

The finding of this research is related to theoretical concepts and a variety of relevant research studies, it can be explained that the TPE assessment instrument which is suspended with a partial scoring can be used to measure the ability of TPE participants in preparing learning implementation plans. Through the partial scoring, this assessment instrument can be used to differentiate and classify the achievement of TPE participants' mastery in preparing learning implementation plans based on their respective abilities. For example, in reaching high category in an item, high ability is totally needed to achieve it. The results of this study clearly indicate the level of difficulty in each category for all items that assess the ability of TPE participants to develop lesson plans. This is in line with the view of Masters and Wright (1997) that the scoring partial credit model is designed to understand the various abilities of participants and distinguish or group participants according to their level of response to items. The results of this research are also related to the view of Verhelst and Verstralen (2008) that item parameters in the partial credit model are interpreted as difficulties in each of the steps that must be completed. The higher the step, the more step or greater ability is needed to achieve it. This means that through the partial credit model used in the TPE assessment model instrument, the assessor can clearly distinguish and classify the capabilities of TPE participants based on their responses at each stage assessed.

This opinion is also supported by Retnawati (2014) who explained that the category score on the partial credit model shows the number of steps to complete the item correctly. Considering that in this study there are four categories, it can be explained that there are four steps in each item to achieve the highest score. Furthermore, Retnawati explained that a higher

ISSN 1822-7864 (Print) ISSN 2538-7111 (Online) https://doi.org/10.33225/pec/19.77.790

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 77, No. 6, 2019

category score requires greater ability to achieve it, and vice versa on a low category score only requires a low ability to also achieve it. If an item follows the partial credit model, then a higher individual ability is expected to have a higher score than an individual who has a low ability.

Therefore, when interpreting the test information function, it is very important to remember the reciprocal relationship between the value of information and the variability of the estimated ability. Baker and Kim (2017) translated information values into standard estimation errors. One thing to note is the square root reciprocity of test information values. Likewise as stated by Hambleton and Swaminathan (1985), one of the most important factors of the test information function is the contribution of each item to the total test information. It is in line with Partchev (2004) who stated if the greater the function information of a test, the smaller the measurement error standard. The greater the information function of a test, the more trusted in measuring the actual test takers' competence. Therefore, the size of the test information function value greatly affects the standard measurement error in estimating the ability of TPE participants with the use of a partial scoring model, because it increasingly provides information on the model chosen in explaining the ability level of TPE participants.

Based on the findings of the information function value and error shows that the value of the information function is very important in estimating the ability of TPE participants in preparing learning plans that are able to provide accurate information on the estimation of the ability of TPE participants followed by the smaller default errors in estimating the ability of TPE participants. This research value information item serves to provide information on the estimation of the ability of TPE participants in accordance with the assessment model chosen by using partial scoring.

Conclusions and Implications

Based on the findings of this research one can indicate that the assessment model with partial scoring accurately estimates the ability of TPE participants in preparing learning implementation plans. These findings can be input for the implementation of follow-up related to the shortcomings and / or weaknesses of TPE participants who are assessed as an effort to improve their ability to prepare learning plans, for example in the assessment of the organizing component of the material. The information is obtained that the material was not written in the form of appropriate points with the formulation of the learning objectives, the material does not contain a brief description of facts, concepts, principles, relevant procedures, and the presentation of the material is not carried out sequentially and systematically. Such information can be input for TPE participants who are assessed to improve it by suggesting that it will improve according to the information obtained. Likewise, for the assessment of other components, the follow-up improvement effort is carried out based on information obtained from the results of the assessment in which parts or stages of each phase, the aspects are judged to be invisible or not conducted by TPE participants in preparing learning implementation plans. Considering that each component assessed in estimating the achievement of TPE participants in preparing a learning implementation plan has a series of stages and a score should be given at each of these stages, so that if the TPE participants' ability is assessed by a partial scoring, the assessment should not only pay attention to the final results, but pay attention to all stages. The stages are to get the mastery achievement of TPE participants in preparing comprehensive learning implementation plans. Besides, the reality on the ground shows that most participant responses are partially correct, true in some steps, but wrong in other steps. For this kind of response, partial scoring is a scoring model that is able to give a description of the limits partially correct in estimating the ability of TPE participants to prepare a learning implementation plan, because this scoring is built with two parameters, namely the ability and level of difficulty of the item and does not require greater threshold from one category to the next. Therefore, it can be

https://doi.org/10.33225/pec/19.77.790 ISSN 1822-7864 (Print) ISSN 2538-7111 (Online)

concluded that the assessment model with partial scoring accurately estimates the ability of TPE participants in preparing learning implementation plans.

Acknowledgements

The authors would like to express great appreciation and the deepest gratitude also would go to all parties involved in this research, especially the participating field supervisors as assessors, Teacher Profession Education Program in the Yogyakarta State University (UNY), Gorontalo State University (UNG), and Islamic State University of Alauddin Makassar (UINAM), Indonesia.

References

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and psychological measurement, 45*(1), 131-142.

Antoro, B. (2017). *Gerakan literasi sekolah dari pucuk hingga akar: sebuah refleksi* [The literacy movement of the school from peak to the root: A reflection]. Jakarta: Direktorat Jenderal Pendidikan Dasar dan Menengah Kementerian Pendidikan dan Kebudayaan.

Baker, F. B., & Kim, S.-H. (2017). *The basics of item response theory using R*. Switzerland: Springer.

Basri, I. Y. (2015). *Kajian tentang program pendididkan profesi guru terhadap eksistensi LPTK* [A study of TPE towards the existence of teacher training university]. Paper presented at the Proceeding FPTK.

Baswedan, A. R. (2014). *Gawat darurat pendidikan di Indonesia* [The emergency of education in Indonesia]. Paper presented at the The Emergency of Indonesian Education. A paper delivered at the meeting between Ministry and Head of Education Offices Indonesia-wide in Jakarta, on December.

Bhakti, C. P., & Maryani, I. (2017). Peran LPTK dalam pengembangan kompetensi pedagogik calon guru [The role of teacher training university in the development of pedagogical competency of prospective teachers]. *Jurnal Pendidikan (Teori dan Praktik), 1*(2), 98-106. doi: 10.26740/jp.v1n2.p98-106

Chang, M. C., Shaeffer, S., Al-Samarrai, S., Ragatz, A. B., De Ree, J., & Stevenson, R. (2013). *Teacher reform in Indonesia: The role of politics and evidence in policy making*. Washington DC: Directions in Development-Human Development World Bank Group.

de Ayala, R. J. (2009). *The theory and practice of item response theory*. New York: Guilford Publications.

Fahdini, R., Mulyadi, E., Suhandani, D., & Julia, J. (2014). Identifikasi kompetensi guru sebagai cerminan profesionalisme tenaga pendidik di Kabupaten Sumedang [The identification of teachers' competency as the reflection of profesionalism of teachers in Sumedang Regency]. *Mimbar Sekolah Dasar, 1*(1), 33-42.

Gareis, C. R., & Grant, L. W. (2014). The efficacy of training cooperating teachers. *Teaching and Teacher Education, 39*, 77-88. doi: https://doi.org/10.1016/j.tate.2013.12.007

Good, T. L. (2008). *21st century education: A reference handbook* (Vol. 1). California: Sage Publications.

Goodwin, A. L. (2010). Globalization and the preparation of quality teachers: Rethinking knowledge domains for teaching. *Teaching Education, 21*(1), 19-32. doi: https://doi.org/10.1080/10476210903466901

Gronlund, N. E., & Linn, R. L. (1965). *Measurement and evaluation in teaching* (Vol. 4). New York: Macmillan.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. New York: Springer Science Business Media, LLC

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publication.

Hidayah, I. (2013). Implementation review of professional education of teachers (PPG) as the implementation of quality management function. *Procedia-Social and Behavioral Sciences, 103*, 467-472. doi: https://doi.org/10.1016/j.sbspro.2013.10.361.

Hotimah, H., & Suyanto, T. (2017). Strategi pendidikan profesi guru (PPG) UNESA dalam mengembangkan kompetensi pedagogik dan profesional peserta PPG Pasca SM-3T [The strategy of TPE program at UNESA in developing pedagogical competency and professionalism of Post SM-3T TPE participants]. *Kajian Moral dan Kewarganegaraan, 5*(01), 242-256.

Hu, B., Qin, L., Sullivan, M., & Templin, J. (2017). Contemporary approaches to psychometrics: item response theory and diagnostic classification models/enfoques contemporáneos sobre psicometría: los modelos de la teoría de respuesta al ítem y los modelos de clasificación de diagnósticos. *Cultura y Educación, 29*(3), 461-491.

Indriyani, S., & Ismandari, D. (2015). Persepsi mahasiswa kependidikan Fakultas Ekonomi Universitas Negeri Yogyakarta terhadap Pendidikan Profesi Guru (PPG) [The perception of economic education students at Yogyakarta State University towards the TPE program]. *Pelita-Jurnal Penelitian Mahasiswa UNY, X*(1), 1-10.

Istiyono, E., Mardapi, D., & Suparno, S. (2014a). *Penerapan partial credit model pada tes pilihan ganda termodifikasi merupakan model alternatif asesmen fisika yang adil* [The implementation of *partial credit model in the modified multiple choice test which is a fair physics alternative assessment model*]. Paper presented at the Prosiding Kongres dan Konferensi Ilmiah Himpunan Evaluasi Pendidikan (HEPI) Tahun 2014, Bali.

Istiyono, E., Mardapi, D., & Suparno, S. (2014b). Pengembangan tes kemampuan berpikir tingkat tinggi fisika (pysthots) peserta didik SMA [The development of the test of physics higher order thinking skill (pysthots) of Senior High School]. *Jurnal Penelitian dan Evaluasi Pendidikan, 18*(1), 1-12.

Krisna, F. N. a. (2017). Alternatif kebijakan permasalahan ketidakmerataan penyebaran guru [Policy alternatives to address teacher distribussion inequality]. *Jurnal Analisis Kebijakan, 1*(2), 1-11.

Mangkunegara, A. A. P., & Puspitasari, M. (2015). Kecerdasan emosi guru, stres kerja, dan kinerja guru SMA [The intelligence of teachers' emotion, work stress, and performance of Senior High School teachers]. *Jurnal Kependidikan: Penelitian Inovasi Pembelajaran, 45*(2), 142-155. doi: 10.21831/jk.v45i2.7491.

Masters, G., & Wright, B. (1997). *The partial credit model. En WJ van der Linden, and Hambleton, RK Handbook of Modern Item Response Theory*. New York: Springer Verlag.

Mizell, H. (2010). *Why professional development matters* (Vol. Learning Forward). Oxford: ERIC.

Ningsih, M. P., Fatchan, A., & Susilo, S. (2016). Program PPG untuk membangun kompetensi guru Geografi (Studi kasus di Universitas Negeri Malang) [TPE program in building competency of Geography teachers (Case study at Malang State University)]. *Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan, 1*(10), 2031-2039.

Nunnally, J. C., Bernstein, Ira H (1994). *Psychometric theory 3E*. New Delhi: Tata McGraw-Hill Education.

Partchev, I. (2004). *A visual guide to item response theory*. Universitat¨ Jena: Friedrich Schiller.

Retnawati, H. (2011). Mengestimasi kemampuan peserta tes uraian Matematika dengan pendekatan teori respons butir dengan penskoran politomus dengan generalized partial credit model [Estimating the ability of mathematic test participants using the approach of item response theory with polytomus score using generalized partial credit model]. *Prosiding Semnas Penelitian Pendidikan dan Penerapan MIPA. UNY*, 53-62.

Retnawati, H. (2014). *Teori respons butir dan penerapannya: untuk peneliti, praktisi pengukuran dan pengujian, mahasiswa pascasarjana* [*Item Response Theory and its implementation: for researcher, practitioner of measurement and testing, students of graduate school*]. Yogyakarta: Nuha Medika.

Retnawati, H. (2016). *Analisis kuantitatif instrumen penelitian* [*Quantitative analysis of research instrument*]. Yogyakarta: Parama Publising.

Retnawati, H., Djidu, H., Apino, E., & Anazifa, R. D. (2018). Teachers' knowledge about higher-order thinking skills and its learning strategy. *Problems of Education in the 21st Century, 76*(2), 216-230.

Retnawati, H., & Munadi, S. (2013). Mengestimasi parameter butir dan kemampuan guru menggunakan model parsial kredit dan parsial kredit tergeneralisasi [Estimating the parameter of item and the ability of teachers in using partial credit model and generalized partial credit]. from Lumbung Pustaka Universitas Negeri Yogyakarta http://eprints.uny.ac.id/id/eprint/22874

Safaruddin, S., Anisa, A., & Saleh, A. (2018). Partial Credit Model (PCM) dalam Penskoran Politomi pada Teori Respon Butir [Partial Credit Model (PCM) in the polytomi scoring of Item Response Theory]. *Jurnal Matematika, Statistika dan Komputasi, 9*(1), 39-48. doi: 10.20956/jmsk.v9i1.3397

Setiajid, S., Susanti, M. H., & Ngabiyanto, N. (2018). *Developing character of nationality participants of SM-3T PPG program through the educational model in Semarang University.* Paper presented at the Annual Civic Education Conference (ACEC 2018).

Sinharay, S., & Haberman, S. J. (2014). How often is the misfit of item response theory models practically significant? *Educational Measurement: Issues and Practice, 33*(1), 23-35. doi: 10.1111/emip.12024

Stronge, J. H. (2018). *Qualities of effective teachers*. Alexandria, VA USA: ASCD.

Templin, J., & Hoffman, L. (2013). Obtaining diagnostic classification model estimates using Mplus. *Educational Measurement: Issues and Practice, 32*(2), 37-50.

Trygstad, P. J., Banilower, E. R., Smith, P. S., & Nelson, C. L. (2014). New Instruments for studying the impacts of science teacher professional development. *Horizon Research, Inc.*, 1-16.

Verhelst, N. D., & Verstralen, H. (2008). Some considerations on the partial credit model. *Psicologica, 29*(2), 229-254.

Wahyudin, D. (2016). Manajemen kurikulum dalam pendidikan profesi guru (Studi kasus di Universitas Pendidikan Indonesia) [Management of Curriculum in TPE program (A case study atEducation University of Indonesia)]. *Jurnal Kependidikan: Penelitian Inovasi Pembelajaran, 46*(2), 259-270.

Widhiarso, W. (2010). *Model politomi dalam teori respons butir* [Polytomi model in the item response theory]. Yogyakarta: Psikologi UGM.

Zamakhsari. (2009). Pendidikan profesi guru: harapan dan tantangan [Teaching Profession Education: Expectation and challenge]. *Mukaddimah, 15*(26), 123-142.

Received: *August 30, 2019* Accepted: *November 29, 2019*

Lian Gafar Otaya *(Corresponding author)*	M.Pd., Student, Doctor Program in Research and Evaluation Education Graduate School of Yogyakarta State University, Jl. Colombo Karangmalang Yogyakarta 55281, Indonesia. Lecturer, Faculty of Tarbiyah and Teaching Training, IAIN Sultan Amai Gorontalo, Jl. Sultan Amai Pone Limboto Barat, 96215, Indonesia E-mail: lian055pasca2016@student.uny.ac.id, lianotaya82@iaingorontalo.ac.id ORCID ID: https://orcid.org/0000-0003-4357-2132
Badrun Kartowagiran	PhD, Professor, Graduate School of Yogyakarta State University, Jl. Colombo Karangmalang Yogyakarta 55281, Indonesia. E-mail: kartowagiran@uny.ac.id ORCID ID:https://orcid.org/0000-0002-8536-5417
Heri Retnawati	PhD, Professor, Mathematics Departement, Mathematics and Science Faculty, Yogyakarta State University, Jl. Colombo Karangmalang Yogyakarta 55281, Indonesia. E-mail: heri_retnawati@uny.ac.id. ORCID ID: http://orcid.org/0000-0002-1792-5873