REiD (Research and Evaluation in Education), 6(2), 2020, 160-173
Available online at: http://journal.uny.ac.id/index.php/reid

Estimating the ability of pre-service and in-service Teacher Profession Education (TPE) participants using Item Response Theory

[*1]Lian Gafar Otaya; [2]Badrun Kartowagiran; [3]Heri Retnawati; [4]Siti Salina Mustakim
[1]Faculty of Tarbiyah and Teacher Training, Institut Agama Islam Negeri Sultan Amai Gorontalo
Jl. Gelatik, Heledulaa, Kota Timur, Kota Gorontalo, Gorontalo 96135, Indonesia
[2]Faculty of Engineering, Universitas Negeri Yogyakarta
Jl. Colombo No. 1, Karangmalang, Depok, Sleman, Yogyakarta 55281, Indonesia
[3]Faculty of Mathematics and Natural Sciences, Universitas Negeri Yogyakarta
Jl. Colombo No. 1, Karangmalang, Depok, Sleman, Yogyakarta 55281, Indonesia
[4]Faculty of Educational Studies, Universiti Putra Malaysia
Persiaran Masjid, 43400 Serdang, Selangor, Malaysia
[*]Corresponding Author. E-mail: lianotaya82@iaingorontalo.ac.id

Submitted: 24 November 2020 | Revised: 31 December 2020 | Accepted: 31 December 2020

Abstract

This research generally aimed to describe the characteristic of the ability of Pre-service and In-service TPE participants using the Item Response Theory, IRT. The research subject comprised 516 participants divided into 239 participants of the Pre-service TPE program and 277 participants of the In-service TPE program using the purposive sampling technique. Data were collected through the technique of observation and documentation. In estimating the item parameter and ability parameter, the IRT model polytomous was implemented, which was furthermore described. This finding shows that the assessor could directly recognize the position of the ability of students in the TPE program based on the item characteristic and the ability between the highest and the lowest grade in the ability scale, so this finding did not only support the implementation of TPE program in Indonesia, but also its applicability was expected to revise the assessment of teachers' performance, the supervision of teachers, field teaching practice, and the assessment in the other teaching fields, so it could be used as an evaluation in revising the assessment model.

Keywords: *teacher professional education, ability, pre-service, in-service, item response theory*

How to cite: Otaya, L., Kartowagiran, B., Retnawati, H., & Mustakim, S. (2020). Estimating the ability of pre-service and in-service Teacher Profession Education (TPE) participants using Item Response Theory. *REID (Research and Evaluation in Education), 6*(2), 160-173. doi:https://doi.org/10.21831/reid.v6i2.36043.

Introduction

Teacher profession education (TPE) program is strongly related to the professionalism of teachers, because it gives the opportunity in mastering knowledge related to teacher profession and gives learning experiences in order to improve the competences of teacher as demanded. Besides, it can enrich the knowledge, theoretical concept, and experience which are deeper in order to be a professsional teacher (Caena, 2011; Galih & Iriani, 2018; Oviyanti, 2016; Petrie & McGee, 2012). It shows the importance of profession education for teachers and it will continuously become an essential issue in improving the high learning quality.

The high learning quality is the key component in the agenda of educational reformation (Hammond & Moore, 2018). Some findings show the quality of education or learning process which really depend on the quality of the teachers (Bahcivan & Cobern, 2016; Gerritsen, Plug, & Webbink, 2017; Kartowagiran, 2012; Le Cornu, 2016; Retnawati, Apino, & Anazifa, 2018). Some studies also found that there is a strong relation between what is done by teachers and the achievement of students. If teachers have good performance, the achievement of students will also be good. Then, the effort to improve the performance of teachers can be done through the evaluation of the quality of teachers (Fahmi, Maulana, & Yusuf, 2011; Steinberg & Garrett, 2016; Stronge, 2018; Sulisworo, Nasir, & Maryani, 2017; Suswantar & Retnawati, 2016). Therefore, it is important to develop the professionalism of teachers in Indonesia.

One kind of profession education conducted in the development of the professionalism of teachers in Indonesia is through The Teacher Profession Education (TPE) program. TPE program is a program from the government which aims to produce teachers/ teacher candidates who are able to master all required competences such as pedagogical competence, professional competence, social competence, and personality competence. This TPE program is expected to produce teachers/teacher candidates who have complete competences such as qualified and characterized besides the other professionalism competences that are required. Besides, TPE program is an absolute requirement for teachers to obtain the experiences that support their professionalism as stated in the national education standard especially to achieve an educator certificate (Amadi, 2013; Anita & Rahman, 2013; Hotimah & Suyanto, 2017; Ningrum, 2012; Nurmaliah, 2018).

TPE program in Indonesia is initiated by the government in order to respond the problems of national education, such as: (1) *shortage*, the lack of teachers especially in the remote and rural area, (2) *unbalanced distribution*, (3) *under qualification*, (4) *low competence* or the incompetent teachers, and also (5) *mismatched*, the irrelevance between the academic qualification and also the course taught (Kemenristekdikti, 2017, 2018). It is also supported with some opinions which state that profession education for teachers can help them to mater the learning materials and can support the readiness to be a professional teacher (Gerdeman, Garrett, & Monahan, 2018; Hotimah & Suyanto, 2017; Robertson, 2017; Wahyudin, 2016). Therefore, to be a professional teacher, it is essential to follow TPE program, even though there are still problems in its implementation. Those problems are referred to the way to improve competence mastery of TPE participants. If there are still many graduates of TPE program who still do not meet the demanded requirement, then the assessment conducted should be questioned whether it has reached the components which can describe the whole competences of TPE participants or not. Then, the research which can estimate the competence mastery of TPE participants by using the item response theory (IRT).

The most important reason to estimate the competence mastery of TPE participants using IRT approach is the assessment estimated using the raw score. It is conducted by summing the scores in every aspect becoming the total score which is divided by maximum score, then the score obtained is compared with the passing grade of TPE program which is 76 (good). This kind of assessment is relative and cannot differentiate students who have good ability, average ability, and low ability based on the component of every aspect assessed by using the classical theory approach. The measurement using classical theory approach has some limitedness such as its real score really depends on the measurement and the testing cannot be compared, because the assessment approach and classical theory approach are random (not systematic), where there is no relation between the real score and the error score. The observation score and the real score change depending on the difficulty level and the scoring, so both of them really depend on the result of students' measured characteristics where the observation score is the only score which can be seen meanwhile the real score and the error score are latent

(Istiyono, Mardapi, & Suparno, 2014; Mardapi, 2017; Retnawati, 2011, 2016; Sumintono & Widhiarso, 2014). It refers to the assessment mostly used in the field where that is unable to obtain accurate information.

Method

This study used a descriptive-explorative approach which aimed to describe the characteristics of the ability of in-service and pre-service TPE participants. Those abilities were the assessment of students' ability in composing lesson plans using the lesson plan instrument and the ability in the implementation of learning by using learning assessment instrument based on the assessment from lecturers in the workshop, field teaching practice or competency test. That assessment then were analyzed to estimate the item parameter from each instrument and the parameter of TPE participants' ability by using IRT polytomous model and the result was described.

Sample

The subject in this research was the groups of TPE program divided into in-service and pre-service TPE program 2019. The program were conducted at three state universities in Indonesia. The total subject of the research was 516 participants comprising 239 participants of pre-service TPE and 277 participants of in-service. The subjects were selected using purposive sampling technique with the consideration that the subjects taken were appropriate with the number of participants from each program.

Instrument and Procedures

Data collected in estimating the ability of the participants of TPE program were divided into two data groups. The first data group is the estimation of the ability of TPE participants in composing lesson plans. The second data group is the estimation of the participants of TPE program in implementing the learning process. The both of data groups were collected through the observation and documentation. The observation technique was used to assess the material mastery of TPE participants which was assessed using lesson plan assessment instrument and learning assessment instrument. Furthermore, the documentation technique was used to assess lesson plans composed by TPE participants in the workshop, field teaching practice or performance practice in the competency test.

The instrument of lesson plan assessment used consisted of 25 items measured by four indicators: the fromulation of competency achievement indicators comprising six items (item 1-6), organizing the materials, methods, media and learning sources comprising six items (item 7-12), organizing the process, assessment and learning evaluation comprising six items (item 13-18), and the implementation of techno pedagogical content knowledge principle comprising seven items (item19-25). Furthermore, the instrument of lesson plan assessment consisted of 20 items measured by four indicators, namely: conducting an educated learning comprising four items (item 1-4), conducting a good learning comprising seven items (item 5-11), facilitating the development of self-potency and characters of participants comprising four items (item 12-15), and also assessing and evaluating the learning comprising five items (item 16-20). Both of the instruments, accurately fulfilled the requirements of validity aspects measured by seven expert judgements refered to Aiken's V table showing all items in the instrument were valid because they fulfilled the required Aiken index which was > 0.75 (Aiken, 1980, 1985). Besides, the estimation of reliability using the inter-rater reliability technique can be seen in Table 1.

Table 1. The Estimation of Inter-rater Reliability

No	Instrument	Criteria	Reliability Coefficient	Explanation
1	Lesson Plan	≥0.70	0.84	Reliable
2	Learning Implementation	≥0.70	0.81	Reliable

Based on Table 1, it is concluded that generally all instruments responded by the rater have had the reliable inter-class coefficient. The instrument was stated as reliable if the coefficient was ≥ 0.70 (Nunnally & Bernstein, 1994).

Data Analysis

The ability of the TPE participants was analyzed using the polytomous item response theory approach and was estimated using the partial credit model (PCM) method through R program with the Extended Rasch Modeling (eRm) package. Furthermore, to describe the assessment result toward the accuracy of TPE participants' ability using the item information function. With that information function, it can state the contribution of the item instrument in revealing the latent trait measured by that instrument and connected with the Standard Error of Measurement. Hambleton et al. (1991) and Retnawati (2014) stated that the information function value had the reversed correlation with SEM, the bigger the information value, the smaller SEM will be or vice versa. Hence, the information function in IRT gave information toward the presumption of the ability level of TPE participants, the smaller the standard error, the more accurate the assessment conducted in predicting the ability. Thus, in this research, the item information function value functioned to provide information toward the presumption of the ability level of TPE participants as the model selected.

Findings and Discussion

The abilities of the in-servive and pre-service TPE participants foccused on this research were the ability in composing lesson plans and the ability in conduction a learning. Both of them were assessed by lecturers in the workshop, field teaching practice and in the competency test through the performance practice. After conducting the assessment toward the ability of TPE participants, it was obtained the ability parameter (θ). The estimation result of the ability of the TPE participants based on the grouping as elaborated as follows.

The Ability of Composing the Lesson Plan

The ability of pre-service and in-service TPE participants in preparing lesson plans is estimated by the partial credit model (PCM) method through the R program with the Extended Rasch Modeling (eRm) package. The results of the analysis are obtained in the form of characteristic items that are completely presented in Table 2.

Table 2. The Result of the Analysis of Item Characteristics in Learning Planning Assessment

Item	Location	Threshold δ_1	Threshold δ_2	Threshold δ_3	Threshold δ_4
A1	0.77	-1.54	0.54	1.33	2.75
A2	0.62	-1.96	0.45	1.54	2.47
A3	0.70	-1.95	0.48	1.34	2.94
A4	0.61	-1.66	-0.03	1.51	2.62
A5	0.61	-1.83	0.29	1.39	2.60
A6	0.79	-1.06	0.19	1.60	2.43
B7	0.73	-1.35	0.23	1.33	2.71
B8	0.60	-1.67	0.02	1.38	2.69
B9	0.70	-1.31	0.15	1.42	2.57
B10	0.67	-1.02	-0.31	1.51	2.52
B11	0.72	-1.36	0.16	1.61	2.50
B12	0.74	-1.06	-0.27	1.47	2.84
C13	0.81	-1.34	0.09	1.64	2.85
C14	0.78	-1.43	0.32	1.37	2.85
C15	0.82	-1.35	0.15	1.47	3.02
C16	0.81	-1.41	0.25	1.49	2.93
C17	0.77	-1.40	0.26	1.41	2.82
C18	0.65	-1.75	0.08	1.54	2.74
D19	0.57	-1.68	0.09	1.27	2.62
D20	0.66	-1.19	0.02	1.22	2.61
D21	0.67	-1.30	0.18	1.26	2.54
D22	0.52	-1.54	-0.06	1.13	2.56
D23	0.51	-1.68	0.19	1.03	2.50
D24	0.49	-1.57	0.08	0.88	2.59
D25	0.65	-1.05	-0.15	1.19	2.63

Based on Table 2, information is obtained that the location parameters of each item vary from 0.49 to 0.82. In addition, the threshold parameter δ_i 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. It was reviewed from the chance of achieving the score, the threshold parameter coefficient δi for each category is different. The higher the achievement category, the higher the threshold coefficient δi. It means that the higher the assessment of lesson plans, the higher the location coefficient, and the more difficult the item with the threshold distribution which is a category of achievement level. The higher the threshold, the more difficult it is to reach the threshold, so participants who have low ability can only reach the threshold (category threshold) too low, participants with medium ability are only able to reach the threshold (category threshold) to intermediate participants with high capability can certainly reach the high threshold category as well. Embretson and Reise (2000) stated that item location reflects the level of ease or difficulty of the item, while the threshold is the threshold between certain categories to be achieved.

Another thing that can be stated based on the results of item analysis is the item characteristic curve. The item characteristic curves are illustrated to make it easier to understand the relationship between each threshold δ i which is the level of difficulty with the participant's ability to reach a certain score or category. The following is an example of the A4 item characteristic curve from the lesson plan assessment assessing the clarity of the formulation of competency achievement indicators using verbs that can be measured or observed. The full results are presented in Figure 1.

Figure 1 is an example of an item characteristic curve from the assessment of lesson plan, that is item 4 evaluates the clarity of the formulation of competency achievement indicators using verbs that can be measured or observed. If related to the results of the item calibration in Table 2, it can be explained that basically item 4 has a location parameter of 0.61 with a threshold parameter δ_1 -1.66, *threshold* δ_2 -0.03, *threshold* δ_3 1.51, and *threshold* δ_4 2.62. Graphically, *threshold* δ_i can be interpreted as the intersection of the curves of each category. From Figure 1, it is clear that to achieve category 2 or to obtain score 2 in item 4, it needs the ability (θ) about -0.03 to 1.51. In addition to the item characteristic curve, another thing that can be explained is the value of the information function. The information function basically can provide maximum information if it is imposed on certain abilities (θ). The following is the result of the value of the information function (IFT) assessment of lesson plan linked to the Standard Error of Measurement (SEM) in Figure 2.

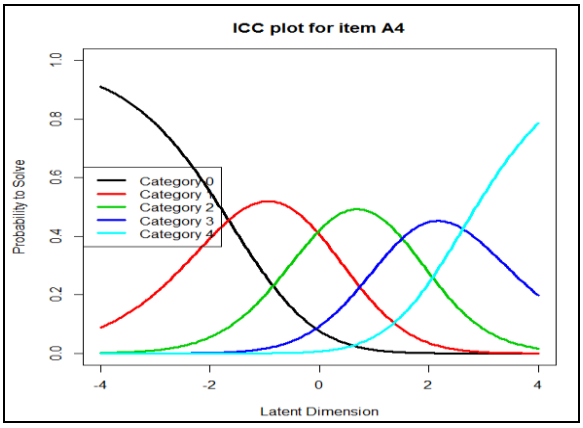

Figure 1. The Curve of Item Characteristic 4 of the Learning Planning Assessment

Figure 2 presents an information function curve from the accumulation of 25 items that assess the ability of TPE participants in preparing the lesson plans. Figure 2 shows a graph of information values (IFT) and measurement errors (SEM) meeting on a capability scale of -4.3 and 0.7. Conversely, when the capability scale is less than -4.3 and more than 0.7, then this instrument has a measurement error greater than the information provided. Another thing from Figure 2 is the instrument information function value of 16.36 on the ability scale (θ) -1.8. Then it can be explained that by knowing the information function value of 16.36, the measurement error coefficient (SEM) obtained by 0.24 indicates the instrument has a higher information value compared to the measurement error. Overall distribution of the estimated results of TPE participants' abilities in compiling a complete learning plan is shown in Figure 3.

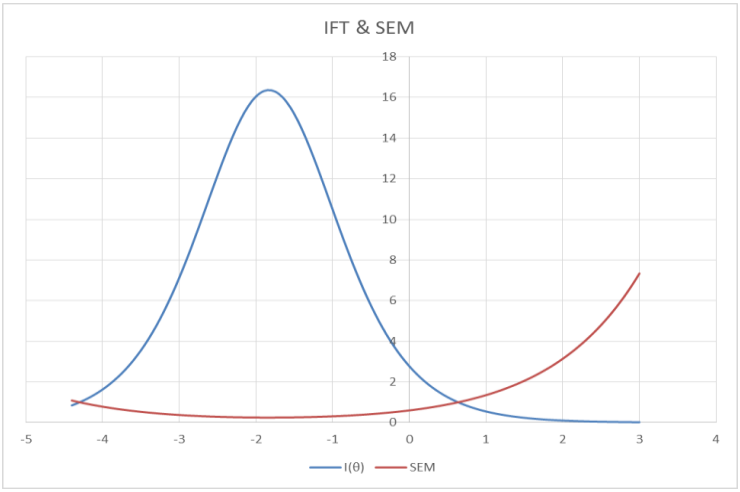

Figure 2. The Converse Relation of IFT and SEM from the Learning Planning Assessment

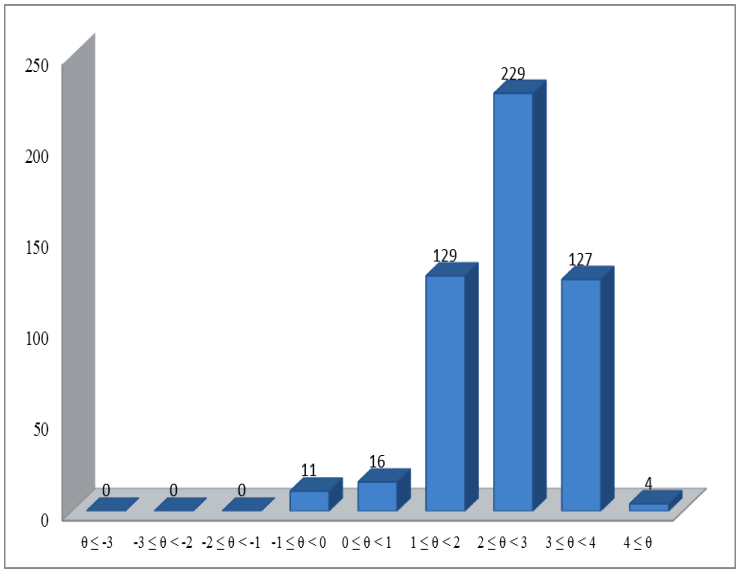

Figure 3. The Distribution of TPE Participants' Ability in Composing Lesson Plans

Figure 3 shows the distribution of the estimated results of TPE participants' abilities in compiling overall learning planning can be said to be good. This is shown by the results of the estimated ability of TPE participants in preparing learning plans dominated by the abilities (θ) 1 to 3. If the results of estimation of TPE Program students' ability in preparing learning plans are grouped based on the ability of pre-service TPE participants and the ability of in-service TPE participants, they are presented in Figure 4 and Figure 5.

Figure 4 shows the estimated results of pre-service TPE participants' ability in developing overall lesson plan can be said to be good, because ideally the expected TPE participants' ability is at least 1 or more. This is shown by the results of the estimated ability of pre-service TPE participants in compiling lesson plans which are dominated by abilities (θ) 1 to 3 where from 239 participants assessed, there were 26% on the ability $1 \leq \theta < 2$, 42% were on ability $2 \leq \theta < 3$, and 29% were on ability $3 \leq \theta < 4$. Besides, Figure 5 shows the estimation of in-service TPE participants' ability which shows the abilities tend to be possessed by participants in the ability (θ) 1 to 3 with the total 277 participants assessed, there were 24% in the ability $1 \leq \theta < 2$, 49% in the ability $2 \leq \theta < 3$, and 21% were on the ability $3 \leq \theta < 4$. The estimation shows that the ability of in-service and pre-service TPE participants in composing the lesson plans are both already good.

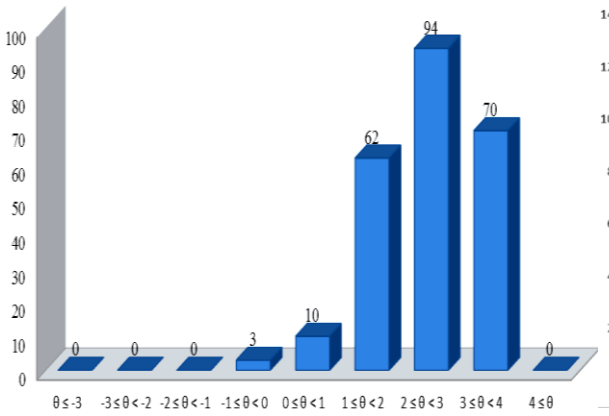

Figure 4. The Distribution of Pre-service TPE Participants' Ability in Composing Lesson Plans

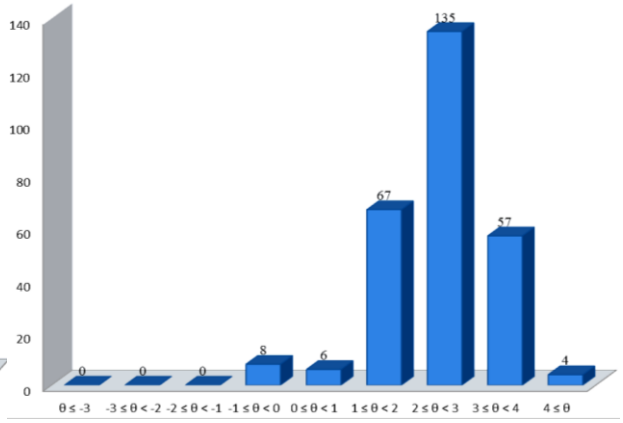

Figure 5. The Distribution of In-service TPE Participants in Composing Lesson Plans

The Ability to Conduct Learning

The second ability which became the analysis unit was the ability of the teacher profession education (TPE) participants in the learning process. Just like in estimating the ability in composing the lesson plans, in this ability was also estimated by the partial credit model (PCM) using the R Program with the help of Extended Rasch Modeling (eRm) package, in which it obtained the analysis of the item characteristics, which was completely shown in Table 3.

Based on the analysis which was served in Table 3, it is obtained information that the location of parameter in every item varied from 0.58 to 0.85. Furthermore, the parameter threshold δi comprised four groups or also known as four intersections. It was the parameter of difficult level of participants in obtaining certain scores when responding item i. As reviewed from the chance of the score achievement, the coefficient of the parameter threshold δi for every category is different. The higher the category of achievement, the higher the coefficient threshold δi. Therefore, it shows that in the assessment of learning process, the higher the coefficient of the location, the harder the item was. The higher the threshold, the more difficult to achieve the threshold, therefore, participants who have the low ability only could achieve the low threshold. Otherwise, participants who had medium ability could only achieve the medium threshold and the participants with high ability could achieve the high threshold. Another thing which could be explained based on the item analysis with the partial credit model (PCM) was the item characteristic curve. The item characteristic curve was described to understand the relation of each threshold δi which was the difficulty level with the ability of participants to achieve certain scores. For example, it can be seen from the item characteristic curve of item C15 about training the students to politely communicate to others and used the appropriate gestures in communication, which are completely presented in Figure 6.

Figure 6 is an example of the item characteristic curve of the learning implementation assessment in item 15. If it is related to the result of item calibration on Table 3, it can be explained that, basically, item 15 has location parameter 0.64 with the parameter threshold δ_1 -1.36, threshold δ_2 -0,13, threshold δ_3 1.29, and also threshold δ_4 2.78. Graphically, the threshold δ_i can be interpreted as the curve intersection of each category. From Figure 6, it can be explained that to achieve category 2 or to obtain score 2 in item 15, it is necessary that the ability (θ) should be around -0.13 to 1.29. In addition to this item characteristic curve, another thing which can be explained is the information function. The result of the IFT is connected to the Standard Error of Measurement (SEM), as completely presented in Figure 7.

Table 3. The Analysis of Item Characteristics of Learning Assessment

Item	Location	Threshold δ_1	Threshold δ_2	Threshold δ_3	Threshold δ_4
A1	0.75	-1.50	0.59	0.95	2.97
A2	0.58	-1.27	-0.10	0.87	2.86
A3	0.65	-1.00	-0.15	0.88	2.91
A4	0.69	-1.17	-0.01	1.16	2.79
B5	0.63	-1.44	0.07	1.03	2.88
B6	0.60	-1.26	-0.22	1.14	2.75
B7	0.73	-1.23	0.04	1.22	2.89
B8	0.71	-1.52	0.14	1.13	3.08
B9	0.70	-1.10	-0.12	1.18	2.85
B10	0.68	-1.07	-0.13	1.07	2.87
B11	0.68	-0.90	-0.37	1.09	2.90
C12	0.62	-1.45	0.19	0.77	2.99
C13	0.58	-1.39	0.10	0.78	2.84
C14	0.64	-1.04	-0.13	0.97	2.77
C15	0.64	-1.36	-0.13	1.29	2.78
D16	0.62	-1.26	-0.19	1.05	2.91
D17	0.65	-1.18	0.02	1.10	2.68
D18	0.67	-1.25	0.19	0.93	2.84
D19	0.64	-1.21	-0.30	1.16	2.95
D20	0.85	-0.91	0.17	1.09	3.07

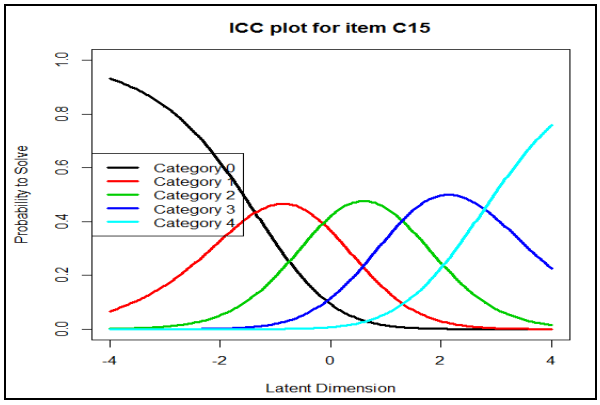

Figure 6. The Distribution of TPE Participants' Ability in Conducting a Learning

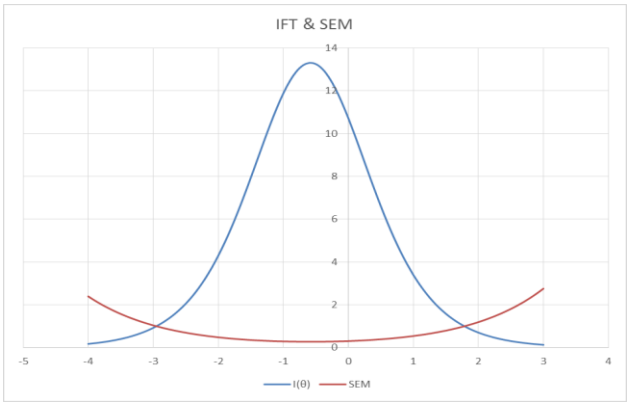

Figure 7. The Converse Relation of IFT and SEM from the Learning Assessment

Figure 7 serves the curve of information function from the accumulation of 20 items in the learning assessment. It shows the graphic of information value and measurement error. Those two function graphics meet in the ability scale -2.9 and 1.8. From both of two abilities, intsrument has higher information value than its measurement eror. Otherwise, when the ability scale is less than -2.9 and more than 1.8, so that assessment has the bigger measurement error than the infromation given. Another thing that can be explained from Figure 7 is the maximum information function value which is 13.3 in the ability scale (θ) -0.6. The bigger the information value, the smaller the SEM will be or vice versa. Therefore, by the identification of the information function value which is 13.3, the SEM obtained is 0.27. The estimation result of TPE participants in the learning process is shown in Figure 8.

Figure 8 presents the estimation of TPE participants in conducting the learning process which is good. It is showed by the estimation of the TPE participants in conducting learning process dominated by the ability (θ) 1 until 3, because the ideal TPE participants' ability that is expected is 1 or more. If the estimation of TPE participants' ability is grouped based on the category of the in-service and pre-service TPE, then it is described in Figure 9 and Figure 10.

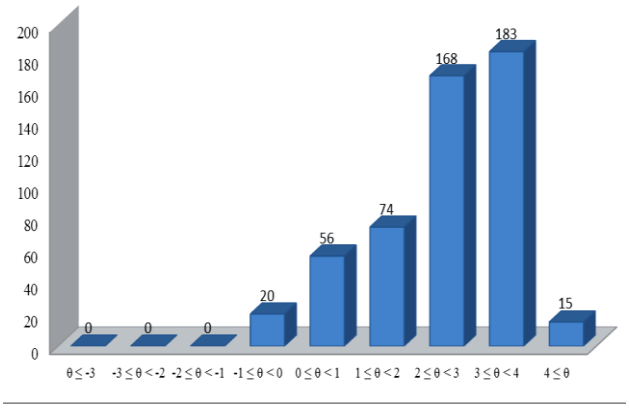

Figure 8. The Distribution of TPE Participants in Conducting the Learning

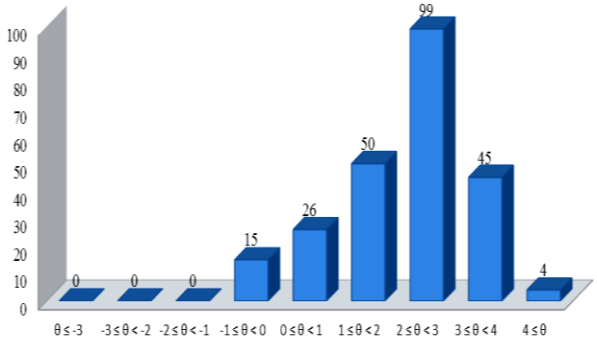

Figure 9. The Distribution of Pre-service TPE Participants' Ability in Conducting the Learning

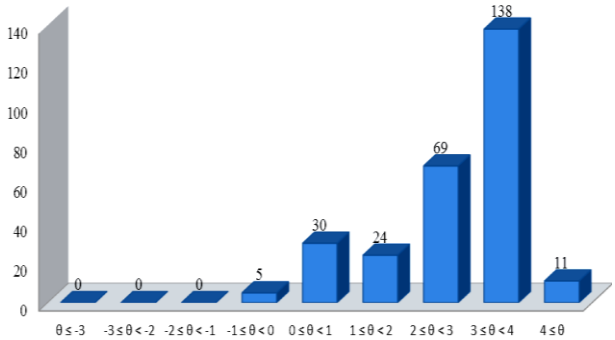

Figure 10. The Distribution of In-service TPE Participants in Conducing the Learning

Figure 9 presents the estimation of in-service TPE participants in conducting learning process which is stated as good, because the ideal TPE partiicpants' ability expected is 1 or more. It is shown by the estimation of pre-service TPE participants' ability in conducting the learning process dominated by the ability (θ) 1 until 3. From 239 participants, there were 21% of participants in the ability $1 \leq \theta < 2$, 41% in the ability $2 \leq \theta < 3$ and 19% in the ability $3 \leq \theta < 4$, while Figure 10 shows the estimation of in-service TPE participants' ability with good results shown by the ability of participants in the ability level (θ) 2 until 3. From 277 participants, there were 25% in the ability level $2 \leq \theta < 3$ and 50% of them were in the ability level $3 \leq \theta < 4$. From the estimation, it can be explained that the ability of in-service TPE participants is higher than the ability of pre-service TPE participants.

One of the findings in this research is that it obtained the assessment or the description of the ability of in-service and pre-service TPE participants using IRT approach. Empirically, the result of the assessment of TPE participants' ability shows the good results, where the ability of participants is dominant in the ability level (θ) 1 until 3. Hence, in the ability of TPE participants from the learning assessment was also dominated by the ability level (θ) 1 until 3. It is supported by Retnawati and Munadi (2013) that the ideal ability parameter is 1 or more. Besides, the ability of participants in composing lesson plans, and conducted the learning process in the in-service TPE participants. This finding was indicated because the participants of in-service TPE program already had teaching experiences rather than the aprticipants of the pre-service TPE program. Dewey (1997) states that experiences are all processes of the living especially when interacting with many things from inside and outside, then that interaction influenced the further interactions. Dewey's point of view became the basis in reflecting the continuous experiences of TPE participants, especially in improving their competences. Paterson (2010) stated education was not only

existed in someone's life, but also as the process which formed the better version of someone. Afterwards, the experiences obtained by the TPE participants were the whole learning processes became the essential experiences for them to be implemented as a professional teacher in the future.

This finding showed that the estimation of pre-service and in-service TPE participants was conducted using IRT approach, so the assessor could directly determine the position of TPE participants' ability in composing lesson plans, becsue IRT approach had assumptions that the latent variable represented by an unidimensional continuum could provide accurate information about the latent attribute or the ability possessed by someone (de Ayala, 2009; Hu et al., 2017). It was also in line with Baker (2001) who stated that one of the aims of IRT was to find the posistion of participants based on the ability scale. Through this information, the assessor could recognize the ability of the participants. Besides, the assessor could also compare the ability among participants in the score determination based on that abiliy scale (θ). This finding also showed the estimation conducted had the high information function value and the small estimations standard error, meaning that the estimation of ability produced was more accurate.

Based on the finding obtained in estimating the ability of TPE participants showed that IRT approach could increase the accuracy of achievement measurement of TPE participants especially in the ability of composing lesson plans. Besides, it was also proved with the accuracy of assessment measured from the information function value and the estimation standard error. It is in line with Baker (2001) who stated that if the parameter can be predicted carefully, so it will be easier to discover the information about the parameter value. It was essential for the assessor to estimate the ability of TPE participants, because the precision which predicted the position of participants' ability depending on the position of someone's ability on the ability scale. Thus, the tendency of the assessment of TPE participants should be directed to the IRT polytomous approach because the ability of TPE participants was ranging as a continuum from the easiest to the most difficult. TPE participants tried to understand or master the expected abilities, so the mastery would be on the position in the continuum, and it was not limited only to the position of the lowest or highest ability. If the ability of the TPE participants is measured by IRT approach, so the measurement results are between the lowest and the highest margin in a continuum.

Through the evaluation of assessment process using IRT approach, it is expected that it can produce the qualified teachers who master the competences and can implement them in the learning process. This effort can realize the improvement of the competences of professional teachers. It is reinforced by some opinions, such as from Biktagirova and Valeeva (2014), Pollard (2014), Liu (2015), and Galih and Iriani (2018) the professionalism of a teacher must be improved not only when teaching in the class, but also before and after the class. Becoming a professional teacher is not only enough with the educator certificate, but a professional teacher should also improve the professionalism continuously, fulfilling the responsibility and duty, conducting self-reflection in making decision to make a better teaching and learning process in the future. Besides, Loughland and Alonzo (2019) state that the criteria of teachers' success in the learning process, really depend on the expectation to fulfil the students' needs. Thus, teachers need to evaluate the learning process as the refection in evaluating and improving self-ability. Teachers have an important role in improving students' critical thinking, improving social and interpersonal communication of students, confidence, learning interest, active participation, and also helping students to prepare themselves as a good citizen. In realizing that, it really depends on the moral imperative of teachers in giving positive response toward the guidance model conducted (Hammond & Moore, 2018; Kuş & Öztürk, 2019). Furthermore, it needs self-awareness from all teachers to always develop their ability to become qualified teachers (Creemers, Kyriakides, & Antoniou, 2012; Gareis & Grant, 2014; Good, 2008; Goodwin, 2010; Rabadi-Raol, 2019; Zhu et al., 2017). As stated by Sheridan and Tindall-Ford (2018),

the assessment of the ability of teachers of teacher candidates becomes more significant to evaluate and improve the learning process to be better in the future.

Conclusion

The finding of this research shows that the TPE participants' mastery is good. It is the absolute requirement to conduct TPE assessment which measures not only the participants' academic mastery, but also the learning achievement and the competency mastery. By recognizing the ability of every TPE participant, mainly the ability in composing lesson plans and the ability in conducting a learning process, the assessment obtained will be more objective, accurate, and accountable in estimating the mastery of TPE participants.

The positive things obtained from the findings are: first, the assessment is designed using IRT polytomous model to determine the level or category achieved by participants based on the response given, so it can collect more information of item characteristic and estimate the TPE participants' ability based on the ability scale. Secondly, it can collect more detail information of TPE participants, it can describe the steps mastered by TPE participants, because the steps assessed from the TPE participants are correct in certain steps, but incorrect in the other steps. Thus, estimating the TPE participants' ability using IRT approach is the choice which possibly gives information of their ability. The higher the parameter of the TPE participants' ability, the bigger the chance they have to do the step by step correctly as the item assessed. Third, the applicability of this assessment is not only used to assess the mastery of TPE participants, but also can be implemented in the assessment of teacher performance, teacher supervision, field teaching practice, and the assessment of other teaching assessments.

References

Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and psychological measurement, 40*(4), 955-959. doi: 10.1177/001316448004000419

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and psychological measurement, 45*(1), 131-142. doi: 10.1177/0013164485451012

Amadi, M. N. (2013). *In-service training and professional development of teachers in Nigeria: Through open and distance education*. Paper presented at the Annual Meeting of the Bulgarian Comparative Education Society.

Anita, N., & Rahman, A. (2013). Penilaian peserta PPG SM-3T prodi PPKN Unesa terhadap pelaksanaan program pendidikan profesi guru (PPG) tahun 2013. *Kajian Moral dan Kewarganegaraan, 3*(1), 409-423.

Bahcivan, E., & Cobern, W. W. (2016). Investigating coherence among Turkish elementary science teachers' teaching belief systems, pedagogical content knowledge and practice. *Australian Journal of Teacher Education (Online), 41*(10), 63-86. doi: 10.14221/ajte.2016v41n10.5

Baker, F. B. (2001). *The basics of item response theory*. Clearinghouse on Assessment and Evaluation.

Biktagirova, G. F., & Valeeva, R. A. (2014). Development of the teachers' pedagogical reflection. *Life Science Journal, 11*(9), 60-63.

Caena, F. (2011). Literature review quality in teachers' continuing professional development. *Education and Training, 20*, 2-20.

Creemers, B., Kyriakides, L., & Antoniou, P. (2012). *Teacher professional development for improving quality of teaching*. Springer Science & Business Media.

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

Dewey, J. (1997). *Experience and education*. Simon & Schuster Inc.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Lawrence Erlbaum Associates.

Fahmi, M., Maulana, A., & Yusuf, A. A. (2011). Teacher certification in Indonesia: A confusion of means and ends. *Center for Economics and Development Studies (CEDS) Padjadjaran University, 3*(1), 1-18.

Galih, A., & Iriani, C. (2018). Persepsi mahasiswa program pendidikan profesi puru (PPG) Pendidikan Sejarah terhadap program PPG. *Jurnal Pendidikan Sejarah, 7*(1), 66-83.

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

Gerdeman, D., Garrett, R., & Monahan, B. (2018). Teacher professional learning through teacher network programs: A multiple case study investigation. *American Institutes for Research*, 1-28.

Gerritsen, S., Plug, E., & Webbink, D. (2017). Teacher quality and student achievement: Evidence from a sample of Dutch twins. *Journal of applied econometrics, 32*(3), 643-660.

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

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

Hammond, L., & Moore, W. M. (2018). Teachers taking up explicit instruction: The impact of a professional development and directive instructional coaching model. *Australian Journal of Teacher Education, 43*(7), 110-133. doi: 10.14221/ajte.2018v43n7.7

Hotimah, H., & Suyanto, T. (2017). Strategi pendidikan profesi guru (PPG) Unesa dalam mengembangkan kompetensi pedagogik dan profesional peserta PPG Pasca SM-3T. *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.

Istiyono, E., Mardapi, D., & Suparno, S. (2014). *Penerapan partial credit model pada tes pilihan ganda termodifikasi merupakan model alternatif asesmen fisika yang adil.* Paper presented at the Prosiding Kongres dan Konferensi Ilmiah Himpunan Evaluasi Pendidikan (HEPI) Tahun 2014, Bali.

Kartowagiran, B. (2012). *Model penilaian kinerja guru.* Paper presented at the Seminar Nasional HEPI Penelitian dan Evaluasi Pendidikan, Pascasarjana Universitas Negeri Yogyakarta.

Kemenristekdikti. (2017). *Pedoman penyelenggaraan pendidikan profesi guru*. Direktorat Jenderal Pembelajaran dan Kemahasiswaan Direktorat Jenderal Kelembagaan.

Kemenristekdikti. (2018). *Pedoman penyelenggaraan pendidikan profesi guru tahun 2018*. Direktorat Jenderal Pembelajaran dan Kemahasiswaan Direktorat Jenderal Kelembagaan.

Kuş, Z., & Öztürk, D. (2019). Social studies teachers' opinions and practices regarding teaching controversial issues. *Australian Journal of Teacher Education, 44*(8), 15-36. doi: 10.14221/ajte.2019v44n8.2

Le Cornu, R. (2016). Professional experience: Learning from the past to build the future. *Asia-Pacific Journal of Teacher*

Education, 44(1), 80-101. doi: 10.1080/1359866X.2015.1102200

Liu, K. (2015). Critical reflection as a framework for transformative learning in teacher education. *Educational Review, 67*(2), 135-157. doi: 10.1080/00131911.2013.839546

Loughland, T., & Alonzo, D. (2019). Teacher adaptive practices: A key factor in teachers' implementation of assessment for learning. *Australian Journal of Teacher Education, 44*(7), 18-30. doi: 10.14221/ajte.2019v44n7.2

Mardapi, D. (2017). *Pengukuran, penilaian, dan evaluasi pendidikan*. Parama Publishing.

Ningrum, E. (2012). Membangun sinergi pendidikan akademik (S1) dan pendidikan profesi guru (PPG). *Jurnal Geografi Gea, 12*(2), 49-55.

Nunnally, J. C., & Bernstein, I. H (1994). *Psychometric theory*. Tata McGraw-Hill Education.

Nurmaliah, C. (2018). *Analisis kemampuan peserta program pendidikan profesi guru (PPG) dalam workshop subject specific pedagogy (SSP) di FKIP Unsyiah*. Paper presented at the Prosiding Seminar Nasional Biotik, Program Studi Pendidikan Biologi Universitas Islam Negeri Ar-Raniry Banda Aceh.

Oviyanti, F. (2016). Tantangan pengembangan pendidikan keguruan di era global. *Nadwa, 7*(2), 267-282.

Paterson, R. W. K. (2010). *Values, education and the adult*. Routledge.

Petrie, K., & McGee, C. (2012). Teacher professional development: Who is the learner? *Australian Journal of Teacher Education, 37*(2), 59-72.

Pollard, A. (2014). *Reflective teaching: In schools*. Bloomsbury Publishing.

Rabadi-Raol, A. (2019). Quality of teacher education and learning: Theory and practice. *Journal of Education for Teaching, 45*(1), 115-117. doi: 10.1080/02607476.2018.1541342

Retnawati, H. (2011). Mengestimasi kemampuan peserta tes uraian Matematika dengan pendekatan teori respons butir dengan penskoran politomus dengan 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*. Nuha Medika.

Retnawati, H. (2016). *Validitas reliabilitas dan karakteristik butir*. Parama Publishing.

Retnawati, H., Apino, E., & Anazifa, R. D. (2018). Impact of character education implementation: A goal-free evaluation. *Problems of Education in the 21st Century, 76*(6), 881-899.

Retnawati, H., & Munadi, S. (2013). *Mengestimasi parameter butir dan kemampuan guru menggunakan model parsial kredit dan parsial kredit tergeneralisasi*. Lumbung Pustaka Universitas Negeri Yogyakarta.

Robertson, S. (2017). *A class act: Changing teachers work, the state, and globalisation*. Routledge.

Sheridan, L., & Tindall-Ford, S. K. (2018). Fitting into the teaching profession: Supervising teachers' judgements during the practicum. *Australian Journal of Teacher Education, 43*(8), 46-64. doi: 10.14221/ajte.2018v43n8.4

Steinberg, M. P., & Garrett, R. (2016). Classroom composition and measured teacher performance: What do teacher observation scores really measure? *Educational Evaluation and Policy Analysis, 38*(2), 293-317. doi: 10.3102/0162373715616249

Stronge, J. H. (2018). *Qualities of effective teachers*. ASCD.

Sulisworo, D., Nasir, R., & Maryani, I. (2017). Identification of teachers' problems in Indonesia on facing global community. *International Journal of Research Studies in*

Education, 6(2), 81-90. doi: 10.5861/ijrse.2016.1519

Sumintono, B., & Widhiarso, W. (2014). *Aplikasi model Rasch untuk penelitian ilmu-ilmu sosial (edisi revisi)*. Trim Komunikata Publishing House.

Suswantar, I. S. D., & Retnawati, H. (2016). Penilaian kinerja guru SMA swasta di Kabupaten Sukoharjo dan faktor-faktor yang mempengaruhi. *Jurnal Evaluasi Pendidikan, 4*(1), 36-44.

Wahyudin, D. (2016). Manajemen kurikulum dalam Pendidikan Profesi Guru (Studi kasus di Universitas Pendidikan Indonesia). *Jurnal Kependidikan: Penelitian Inovasi Pembelajaran, 46*(2), 259-270.

Zhu, X., Goodwin, A. L., & Zhang, H. (2017). *Quality of teacher education and learning: Theory and practice*. Springer Nature.