

PAPER • OPEN ACCESS

Differential item functional analysis on pedagogic and content knowledge (PCK) questionnaire for Indonesian teachers using RASCH model

To cite this article: B D Rahmani 2018 *J. Phys.: Conf. Ser.* **948** 012061

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the [collection](#) - download the first chapter of every title for free.

Differential item functional analysis on pedagogic and content knowledge (PCK) questionnaire for Indonesian teachers using RASCH model

B D Rahmani

Universitas Muhammadiyah Prof. DR. HAMKA

E-mail: bitadwi@uhamka.ac.id

Abstract. The purpose of this paper is to evaluate Indonesian senior high school teacher's pedagogical content knowledge also their perception toward curriculum changing in West Java Indonesia. The data used in this study were derived from a questionnaire survey conducted among teachers in Bandung, West Java. A total of 61 usable responses were collected. The Differential Item Functioning (DIFF) was used to analyze the data whether the item had a difference or not toward gender, education background also on school location. However, the result showed that there was no any significant difference on gender and school location toward the item response but educational background. As a conclusion, the teacher's educational background influence on giving the response to the questionnaire. Therefore, it is suggested in the future to construct the items on the questionnaire which is coped the differences of the participant particularly the educational background.

1. Introduction

Teacher Competency Test or *Ujian Kompetensi Guru (UKG)* is a test to identify teacher's competencies in pedagogic and content knowledge of a subject matter mastered by each teachers in Indonesia. Since 2012, Indonesia Ministry of Education Board has been administering the *UKG* as the prerequisite of being a certified-teacher. However, since 2015, *UKG* has been conducted to all teachers in all over Indonesia which covers 192 subject matter/skills as a parameter in identifying the level of teacher's competence both in professional and pedagogic competence. Unfortunately, *UKG* in 2015 did not run successfully since there were many problems faced on the test such as the construct of item test which was not in line with the material, the difficulty level of the test and even the technical problem on the internet connection and many more. Due to these problems, *UKG* result could not become the most single accurate sources of leveling the teacher's competencies. There should be another instrument to support the result of *UKG*. The teacher self-assessment was seen as a good alternative answer of the problem. The teacher self-assessment test allows the teacher to reflect his/her self on teaching learning process. Moreover, the self-assessment is a powerful technique for self-improvement efficacy in teaching [1]. Thus, this project aims to design a self-assessment questionnaire on pedagogic, content and professional knowledge of the teacher. Despite the differences of sample in filling the questionnaire, it should have been checked whether all the items on the questionnaire were functioning well or not to differentiate the heterogeneity of sample from their gender, educational background also on school location. Thus, the differential item analysis was conducted to analyze the data.

Despite the importance of having a valid and reliable item on the questionnaire, this study has been undertaken to answer some of the following questions: (1) Are there any differences on gender, educational background and/or school location in responding the item on the questionnaire of PCK (Pedagogical and Content Knowledge) and Curriculum Change?; and (2) If yes, how big are the differences on each dimensions?



Due to the limitation of the research, this research only focuses on two aims since those purposes are the basic prerequisite for further analysis. Those two aims are: (1) to investigate whether there are any differences in gender, educational backgrounds and or school location in responding the questionnaire of PCK (Pedagogical and Content Knowledge) and Curriculum Change.

To examine the differences of each item on questionnaire for PCK (Pedagogical and Content Knowledge) and Curriculum toward gender, educational background and or school location. In Indonesia, a ministry of education through *Undang Undang Guru dan Dosen* 2005 or the 2005 Teacher and Lecturer Policy managed that the teachers should have pedagogical, personal, social and professional competences. In order to have all those competencies, the Indonesian Ministry of Education had already ensured the opportunity to improve teachers' quality and competencies through professional development activities. Professional development is an ongoing process of teachers' training, aiming to support teachers' efforts to understand and form their teaching practices [2]. However, the teacher professional program in Indonesia was run based on the teacher competencies level. The teacher's competencies its self was been mapping by the teacher competency test administered by Ministry of Education in 2015. However, the result of UKG was not satisfying. It happened since the curriculum changing also was influenced on teacher pedagogic and content knowledge.

Therefore, there should be another parameter to map the teacher's competencies in Indonesia. Teacher self-assessment is one of the solution in facing the problem. By doing the self-assessment, the teacher could reflect what had already been happened in the classroom and it would be expected to increase the teacher's awareness and efficacy [1]. Moreover, the teacher self-assessment could be as another supporting evidence to mapping the teacher's competencies in Pedagogical and content knowledge aspects.

Many teacher self-assessments has been developed recently. One of them was developed by Schmidt et al [3] Sahin [4]. These self-assessments were more focusing on pedagogical and content knowledge (PCK) of the teacher. It is expected that the teacher would identify by themselves whether they had already implement PCK on their teaching practice. Moreover, due to the happening issue particularly about curriculum change in Indonesia, this questionnaire had already been adjusted to cover that problem also. However, the new construction of the questionnaire really needed a further investigation. The questionnaire had been distributed to several teachers with different gender and educational backgrounds in several schools across Indonesia. Seeing the differences on the sample, the differential item functioning analysis (DIFF) should be conducted on this research. Moreover, DIF can be defined as a statistical phenomenon that occurs when "two individuals with equal ability but from different groups do not have equal probability of success on the item" [5]. In other words, DIF occurs when examinees from different groups show differing probabilities of success on (or endorsing of) the item after matching on the underlying ability that the item is intended to measure. DIF analysis, therefore, is a statistical tool that can be used to improve how a test behaves across groups and to reduce group-based differentials that are not relevant to the construct being measured. In terms of utility, DIF analysis is mainly a tool to assess test fairness, investigate threats to validity, and explore the underlying processes in item responses across groups [6].

2. Methodology

This study gathered a primer data analysis of self-assessment questionnaire that used to investigate the Indonesian ESL teachers based on their Pedagogical and Content Knowledge (PCK). This research is aimed at examining the item response differences toward gender, educational backgrounds also on school locations. One parameter (Rasch model) was used on gathering the data. An instrument was used

in this study is in the form of self-assessment questionnaire adapted from Schmidt et al [3] and Sahin [4]. The instrument consists of 25 items for measuring Indonesian ESL teacher knowledge of two categories/ factors which include Pedagogical Content Knowledge (PCK) and Curriculum Changing Knowledge (CC). The 25 questions on the questionnaire used Likert Scale. The Likert Scale can be used to collect attitude data [7] including the human condition [8]. The participants of this study were sixty one Indonesian English Teacher (ESL), most of whom are from West Java, some from East Java and Jakarta, Indonesia. The data collection was held on National Gathering of Senior High School English Teachers 2017 in Bandung, West Java on July 6th 2017. The population was about 200 teachers on the list. However, the questionnaire was collected only from 61 respondents as the sample of this research. In addition, Differential Item Functional Analysis (DIFF) was used to conduct the analysis of the differences of each item response toward the differences of gender, educational background and school location on the questionnaire by using a software Conquest 4.0. The Conquest produced a significant value, estimate and error score. The significant value was counted firstly and it should be less than 0.05 to indicate the items has a difference on each category. Next, when the significant value indicated the difference on each category, the magnitude also (T value) should be counted secondly to see the deviance on each item toward the same category.

3. Results and Discussion

The first running of the data by using Conquest 4.0 in analyzing the differences between item response of Category 1 (PCK) to Gender (Male, Female) showed that there were not any significant differences. It is shown by Chi-square test of parameter equality = 20.41, df = 13, Sig Level = 0.085. In short, male and female teachers were given the same trait in responding the category 1 (PCK) on the questionnaire. Then, the next calculation was counting the differences between items toward the educational background of the teacher. There were two categories of teacher's qualification namely, bachelor degree and master degree. The result showed that there was a significant difference on two categories in responding the item on factor 1 (PCK). It can be seen on the following table below.

Table 1. The Diff of Factor 1 (PCK) on Educational Background

Item	Bachelor Degree	Master Degree	Magnitude	T
1	0.529	-0.529	1.058	2.213953
2	-0.022	0.022	-0.044	-0.09752
3	0.431	-0.431	0.862	1.803806
4	-0.944	0.944	-1.888	-2.24519
5	0.158	-0.158	0.316	0.700344
6	0.038	-0.038	0.076	0.168437
7	0.369	-0.369	0.738	1.565892
8	1.199	-1.199	2.398	3.36734
9	0.37	-0.37	0.74	1.570135
10	-0.655	0.655	-1.31	-2.66
11	0.212	-0.212	0.424	0.915787
12	-0.209	0.209	-0.418	-0.66939
13	-0.533	0.533	-1.066	-1.48474
14	-0.944	0.944	-1.888	-2.65996

Chi-square test of parameter equality = 20.41, df = 13, Sig Level = 0.085

The Table 1 shows that the teacher with bachelor degree could answer easily for six numbers (symbolized by negative value i.e item number 2, 4, 10, 12, 13 and 14). Meanwhile, the most difficult for them was number (1) *I can use teaching methods and techniques that are appropriate for a learning environment* (symbolized with the biggest positive value, 0.529). It happened since the question was asking about kinds of teaching methods and techniques used by them in the classroom.

The limitation of the knowledge on kinds of teaching and techniques made the teacher answer unconfidently on their competence. In contrast, the teacher with master degree could answer easily most of the items particularly number 1, 3, 5, 6, 7, 8, and 11. However, the most difficult item for them was item number (14). *I can use appropriate teaching methods and techniques to support students in developing their language skills*. This question was asking not only the appropriate teaching methods and techniques but also how the teaching methods and techniques supported the development of student's language skills. This questions made the teacher think for twice to choose their competence scale whether their competence had already developed the student's skill or not.

In addition, the Table revealed the fact that there was differences between the teacher educational backgrounds. It can be seen on item estimate difference greater than 0.5 logits as the criterion for detecting DIF in 14 items on category 1 [9]. Tennant and Pallant [10] stated that to determine whether there is a difference, three indicators were used, to wit: (1) t value is > 2.0 ; (2) DIF contrast value is > 0.5 ; and (3) p (Probability) value is < 0.05 .

An item needs to meet those three conditions to be considered bias and be dropped from the instrument. However, if the item meets only one of the conditions, it should not be dropped but it should be separated and fixed. Those are item numbers 1, 4, 8, 10, and 14 which had the magnitude score above 0.5 logits and the T value is greater than 2. In short, there were 5 items on category 1 which has a significant difference toward the educational background of the participant. However, the five items should not be dropped since the probability of the items was significant which is more than 0.05 ($p=0.085$). In addition, those five items were bias on educational background because of the following reasons below

Item 1: I can use teaching methods and techniques that are appropriate for a learning Environment

The magnitude value of this item was 1.058 (.0.5) and the T value was 2.213953 (>2). The item was biased on teacher's educational background since there was the words "*teaching methods and techniques*" on the questions which asked the teacher's understanding and knowledge to answer the question. In this case, the teacher with master degree were performing well. It is seen from the negative value (-0.529) which means the item was quite easy for them.

Item 4: I can collaborate with school stakeholders (students, parents, teachers, etc.) to support student's learning.

The magnitude value of this item was -1.888 (.0.5) and the T value was -2.24519 (>2). The item was biased on teacher's educational background since there was the words "*collaborate with school stakeholders*" on the questions which asked the teacher's experience engaging the stakeholders involved to support the student's learning. Interestingly, the teacher who graduated from bachelor degree were performing well on this question indicated from the negative value (-0.944) which means the item was quite easy for them. On the other words, the teacher with bachelor degree tended to have a better networking and experience to the stakeholders to support student's learning rather than the teacher with master degree.

Item 8: I can express my ideas and feelings by writing in English

The magnitude value of this item was 2.398 (>0.5) and the T value was 3.36734 (>2). The item was biased on teacher's educational background since the item was asking the teacher's ability on writing. In fact, the teachers who graduated from master degree were doing better in responding this item indicated from the negative value (-1.199) which means the item was quite easy for them. On the other words, the educational background was in line with teacher's competencies.

Item 10: I can understand texts written in English

The magnitude value of this item was -1.31 (>0.5) and the T value was -2.66 (>2). The item was biased on teacher's educational background since the item was asking the teacher's understanding on English text. Interestingly, this item was responded successfully by the teacher with bachelor degree indicated by the negative value (-0.655) which means the item was quite easy for them.

Item 14: I can use appropriate teaching methods and techniques to support students in developing their language skills

The magnitude value of this item was -1.888 (>0.5) and the T value was -2.65996 (>2). The item was biased on teacher's educational background since the item was asking the teacher's knowledge on teaching methods and techniques to support the student's language skills. Interestingly, this item was responded successfully by the teacher with bachelor degree indicated by the negative value (-0.944) which means the item was quite easy for them. It happened since the teacher with master degree were more carefully in responding this item since they had learnt so many teaching methods and techniques on language skills.

The differences on educational background between bachelor and master teacher also happened on category 2 (Curriculum Change perception) of teacher. It can be seen on the following table

Table 2. The Diff of Factor 2 (CC) on Educational Background

Item	Bachelor Degree	Master Degree	magnitude	T
16	0.327	-0.327	0.654	1.418552
17	-0.41	0.41	-0.82	-2.87043
18	-0.492	0.492	-0.984	-3.41075
19	0.406	-0.406	0.812	1.76126
20	0.6	-0.6	1.2	2.274874
21	0.201	-0.201	0.402	0.938142
22	-0.363	0.363	-0.726	-2.55403
23	0.189	-0.189	0.378	0.882133
24	-0.595	0.595	-1.19	-4.24978
25	0.136	-0.136	0.272	0.902972

Chi-square test of parameter equality = 28.15, df = 9, Sig Level = 0.001

The Table 2 shows that the teacher with bachelor degree could answer easily for 4 numbers (symbolized by negative value i.e. item number 17, 18, 22 and 24. Meanwhile, the most difficult for them was number (20) (*I know how to implement the curriculum 2013 in my class* - symbolized with the biggest positive value, 0.6). It happened since the question was asking abstractly how to run the curriculum 2013. The major problem of the question was on which part of the curriculum that the question was asked for. The limitation of teacher's knowledge on curriculum practice made the teacher who graduated from bachelor degree could not be confident enough in responding this item. On the other hand, the teacher with master degree could answer easily on the items particularly number 16, 19, 21, 23, and 25. However, the most difficult item for them was item number "*Curriculum 2013 is easy to implement in my class*". This item was seen as the unclear statement since there was not any precise word to define in which part of curriculum 2013 to be implemented. In addition, the Table 2 also showed the fact that there was differences between the teacher educational backgrounds on category 2 (curriculum change). It can be seen on item estimate difference greater than 0.5 logits as the criterion for detecting DIF in 10 items on factor 2. The same criteria was used to analyse the item on factor 2. As a result, there are five items (item number 7, 18, 20, 22, and 24) which has the magnitude score above 0.5 logits and the T value is greater than 2. In short, there were 5 items on factor 2 which has a significant difference on educational background toward the educational background of the participant. However, the five items should not be dropped since the probability of the items was quite significant 0.01 ($p=0.01$). In addition, those five items were bias on educational background because of the following reasons below.

Item 17: I know the reason of curriculum changing

The magnitude value of this item was -0.82 ($>-/+ .0.5$) and the T value was -2.87043 (>2). The item was biased since the question was asking about the teacher's knowledge on curriculum changing. In this case, the teacher with master degree were performing well indicated from the negative value (-0.327) which means the item was quite easy for them.

Item 18: I know the purpose of new curriculum in teaching

The magnitude value of this item was -0.984 (<0.5) and the T value was -3.41075 (>2). The item was biased on teacher's educational background since it was asking about teacher's knowledge on curriculum change. Interestingly, the teacher with bachelor degree performed well than the master degree research in giving the response on this item indicated by negative value (-0.492). This result was contradictory than the previous item number (item 17) which was also asking about the knowledge of the curriculum changing.

Item 20: I know how to implement the curriculum 2013 in my class

The magnitude value of this item was 1.2 (>0.5) and the T value was 2.274874 (>2). The item was biased since the question was asking about the teacher's knowledge on curriculum changing. In this case, the teacher with master degree were performing well indicated from the negative value (-0.6) which means the item was quite easy for them.

Item 22: Using Curriculum 2013 helps the teaching process

The magnitude value of this item was -0.726 (>0.5) and the T value was -4.24978 (>2). The item was biased on teacher's educational background since it was asking about teacher's knowledge on curriculum change. Interestingly, the teacher with bachelor degree performed well than those with master degree in giving the response on this item indicated by negative value (-0.363) which means the item was quite easy for them.

Item 24: (24). Curriculum 2013 is easy to implement in my class

The magnitude value of this item was -1.19 (>0.5) and the T value was -4.24978 (>2). The item was biased on teacher's educational background since it was asking about the implementation of the the new curriculum in the classroom. Surprisingly, the teacher with bachelor degree performed well than those with master degree in giving the response on this item indicated by negative value (-0.595) which means the item was quite easy for them.

4. Conclusion

Differential Item Functioning (DIFF) had been conducted to analyze the differences on gender, education background, and school location on questionnaire of Pedagogical and Content Knowledge (PCK) and Curriculum Changing (CC) for Indonesian ESL Senior High School Teacher. The questionnaire comprised of 25 item in which 15 items related to PCK and the rest were about Curriculum Changing. In addition, due to there were two categories on the questionnaire, therefore the questionnaire was categorized into two factors namely Pedagogical Content Knowledge (PCK) as factor 1 and Curriculum Change (CC) as factor 2. The analysis on gender, background education and school location were conducted on each category. In short, the result on analyzing the first factor revealed the fact that there was no any difference on gender and school location but educational background among the teachers with bachelor degree and master degree in giving the responses on the questionnaire. It could be observed by a magnitude value which was more than 0.5 and T value which was more than 2.0. In fact, there were 5 items which met those criteria. However, those items were still be maintained since the significant value of all items were more than 0.05 ($p = 0.085$) which meant no significant to be deleted. On the other hand, the result on analyzing factor 2 showed that the differences also were based on educational background instead of gender and school location. The significant value on educational background of factor 2 was less than 0.05 ($p = 0.01$) which indicated that there was a significant differences between teacher with bachelor and master degree in responding the item. In fact, it had been also found that there were also five items which had been identified. It meant that there was a different value between teacher with bachelor and master degree. As a conclusion, the questionnaire had different responses where distributed to Indonesian ESL teachers in senior high school particularly on their education background rather than on gender and school location. It is been suggested to do a refinement of the questionnaire which the items could cope all level of education among the teachers' education background.

References

- [1] J. A. Ross and C. D. Bruce 2007 Teacher self-assessment: A mechanism for facilitating professional growth *Teach. Teach. Educ.* 23 **2** pp. 146–159
- [2] L. Darling-Hammond 1994 *Professional development schools: Schools for developing a profession* ERIC
- [3] D. A. Schmidt, E. Baran, A. D. Thompson, P. Mishra, M. J. Koehler, and T. S. Shin 2010 Technological Pedagogical Content Knowledge (TPACK): The development and validation of an assessment instrument for preservice teachers *Journal of Research on Technology in Education* 42 **2**. pp. 123–149
- [4] I. Sahin 2011 Development of survey of technological pedagogical and content knowledge (TPACK),” *TOJET Turkish Online J. Educ. Technol* 10 **1**
- [5] L. Shepard, G. Camilli, and M. Averill 1981 Comparison of Procedures for Detecting Test-Item Bias with Both Internal and External Ability Criteria *J. Educ. Stat.* 6 **4** pp 317–375
- [6] B. D. Zumbo and M. N. Gelin 2005 A Matter of Test Bias in Educational Policy Research: Bringing the Context into Picture by Investigating Sociological/Community Moderated (or Mediated) Test and Item Bias *J. Educ. Res. Policy Stud.* 5 **1** pp. 1–23
- [7] R. Likert 1932 A technique for the measurement of attitudes *Archives of Psychology* 22 p. 55
- [8] S. Hales 1986 Rethinking the business of psychology *J. Theory Soc. Behav.* 16 **1** pp 57–76
- [9] J. D. Scheuneman and R. G. Subhiyah 1998 Evidence for the validity of a Rasch model technique for identifying differential item functioning *J. Outcome Meas.* 2 **1** pp. 33–42
- [10] J. F. Pallant and A. Tennant 2007 An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS) *Br. J. Clin. Psychol.* 46 **1** pp 1–18