Multirater Assessment to Teacher Professionalism based on Pedagogical Content Knowledge

Maryati*
Yogyakarta State University, Indonesia

*Email: maryati@uny.ac.id

Abstract. The Indonesian Government has performed teacher certification program since the establishment of the Teacher Law No. 14 year 2005 and PCK has been acknowledged as the indicator of science teachers’ professionalism in some countries. Therefore, this research aims to explore teachers’ PCK as a case study to 20 middle school science teachers with different teaching experiences in the special Region of Yogyakarta, Indonesia. This research was conducted in PCK planning and enacted PCK by collecting lesson plans, recorded the learning activities in the classroom, and structured interviews. The assessment to all research material was employed by 6 raters using PCK assessment rubric consisting of 12 items that was later analyzed with Many Facet Rasch Model. The analysis result informed that all items had good validity and reliability. This study found that only 15% of participants were classified as good teachers’ PCK. In-depth analysis on the difficulty level item concluded that the teacher participant’s PCK ability on the representation and conceptual teaching strategies aspects still needed to be improved.

Keywords: Multirater assessment; Pedagogical content knowledge; Rasch model.

1. Introduction
Many programs and rules to enhance teacher’s quality have been employed by Indonesian Government since their Independence [1]. Teachers’ certification is the last on-going program by the settlement of the Teacher Law No. 14 of year 2005 [2]. Concerning with the teacher’s certification program, Indonesian Government has settled 4 competence standards for professional teachers (Teacher Law No. 14 of year 2005), they are: (1) profession competence; (2) pedagogical competence; (3) social competence; and (4) personal competence. Meanwhile, the indicator of teacher professionalism that has been accepted by some countries in the world is based on the pedagogical content knowledge (PCK) [3], [4] [5], [6]. PCK is defined as “the blending of content and pedagogy” referring to the understanding of “how particular topics, problems or issues are organized, represented and adapted to the diverse interest and abilities of learners and presented for instructions” [4]. Hence, Van Driel &Berry (2012) [7] highlight that the development of teacher profession should focus on PCK although PCK is very complex.

As the indicator of teachers’ professionalism, PCK refers to the particular topic because every topic, especially in science education has its own characteristics that support the application of specific strategies in its learning [8], [9]. Hence, the content for teaching is in topic level, not discipline or subject [10]. Hence, Topic-Specific Pedagogical Content Knowledge (TSPCK) construct to assess teachers’ PCK quantitatively [11]. This construct adopts five PCK components, namely: Learner prior
knowledge; Curricular saliency; What makes a topic easy or difficult to understand; Representations; and Conceptual teaching strategies [12].

This article is going to discuss about middle school science teachers’ PCK of matter characteristic, which is a case study to 20 middle school science teachers with different teaching experiences, especially in the special Region of Yogyakarta, Indonesia. This topic is the key component in some science curriculums in middle school [13]. According to some researchers, some of the learning strategies that can enhance the understanding to matter characteristic learning are the use of real objects and computer animation [14] and role playing demonstration [15]. The PCK assessment is conducted by involving multi-rater and the result is analyzed with Many Facet Rasch Model (MFRM). The analysis result is going to be used to map the PCK ability on 4 PCK components according to [12]. Since PCK develops with the teaching experiences [16], therefore these are some of the research questions to conduct this research:

a. What is the PCK ability level of 20 teacher participants of this research like?
b. What is the positive and negative aspects of middle school science teachers’ PCK ability in this research sample, especially in teaching matter characteristic topic?

The rest of this paper is organized as follow: Section 2 describes the proposed research method. Section 3 presents the obtained results and following by discussion. Finally Section 4 concludes this work.

2. Material and Proposed Method

This is a quantitative study as case study of Indonesian science teacher in a natural setting [17].

2.1. Data

The participants (sample) in this research consisted of 20 teachers working voluntarily with teaching experiences as shown on Table 1.

| Teacher experiences (year) | Number | Teacher initials |
|---------------------------|--------|-----------------|
| 0-5                       | 10     | 1,2,6,11,12,9,14,15,16,17 |
| 15-10                     | 1      | 13              |
| 10-15                     | 3      | 8,10, 19       |
| 15-20                     | 3      | 7,5,18         |
| >20                       | 3      | 3,4,20         |

The research material was the lesson plan to teach the topic of “matter characteristic”, teaching activities recordings, and interview result recordings before and after the observation from all teacher participants involved in this particular research. The assessment was employed by 6 raters using PCK assessment rubric consisting of 12 items (see Table 2). In the assessment process, every teacher was assessed by 3 raters by scoring 1-4 (limited to exemplary) in accordance with the fact on the lesson plan, teaching recording, and interview result. Although this quantitative research only involves 20 samples, involving multi-raters will obtained data with a number of 20 x 12 x 6 = 1440 (if no data was lost).

2.2. Instrument

This study employed a rubric consisting of 12 items to assess teachers’ PCK integrating the specific content knowledge of matter characteristic into 4 components proposed by Geddis et al. in [12].
Table 2. The Descriptions of Rubric Items of PCK Assessment with PNM Topics

| PCK Components                  | Codes of Items | Source                  |
|---------------------------------|----------------|-------------------------|
| Curricular saliency            | C01; C02; C03  | Lesson plans & interviews|
| Learner’s Prior Knowledge      | P04            | Lesson plans & interviews|
| Level of difficulty            | D05            | Lesson plans & interviews|
| Representation & Conceptual    | S06; S07; S08; S09; S10; S11 & S12 | Lesson plans, Interviews, Video recordings |
| teaching strategies            |                |                         |

Each item in the rubric is rated on a four-point scale (1=limited, 2= basic, 3=proficient, and 4=exemplary).

2.3. Data analysis Using many Facet Rasch Model (MFRM)

The raw data in scores (1 – 4) from all raters were submitted and recorded in Microsoft Excel, and then analyzed using Many-Facet Rasch Model (MFRM) using the Facets 3.7.2 software. MFRM would analyze the science teachers’ PCK, item difficulties, and rater severity at the same time and display them in one diagram Wright that was easy to understand. The MFRM calibration is based on the severity \( (C) \) of the raters \( (j) \); raters’ severity is considered as an estimate of the probability of teachers’ PCK \( (n) \) on rubric items \( (i) \) to determine the threshold category \( (k) \) for the raters \( (j) \) with the following equation (1):

\[
P_{\text{nikj}} = \frac{e^{(\beta_n - \delta_i - F_k - C_j)}}{1 + e^{(\beta_n - \delta_i - F_k - C_j)}}
\]

This equation involves teachers’ PCK \( (\beta_n) \), rubric items’ difficulty level \( (\delta_i) \), the difficulty level threshold \( (F_k) \), and raters’ severity \( (C_j) \) [18], [9].

The MFRM is an accurate analysis to teachers’ performance with small sample. The information that could be obtained from MFRM was raters’ accuracy, reliability, and validity based on the value of outfit means-square (MNSQ), outfit z-standard (ZSTD) and point measure correlation (Pt Mea Corr) [18], [20].

3. Result and Discussion

This section presents the obtained results and following by discussion.

3.1. The PCK ability of teacher participant

The Table 3 below is the statistics summary that showed good reliability for 3 facets, they are ratee, item, and rater [20]. High item reliability showed that item had measurement that was reproducible. Index separation was useful to differentiate items that were easy and difficult, while index separation person (both rate & rater) showed how good the instrument used to group ratee and rater.

Table 3. Statistics summary of teachers’ PCK analysis with MFRM

| Facet   | Mean | SD | Separation | Reliability | Outfit MNSQ | ZSTD | Pt. Mea |
|---------|------|----|------------|-------------|-------------|------|---------|
| Ratee   | 0.00 | 0.80 | 2.50 | 0.90 | - | - | - |
| Item    | 0.00 | 0.96 | 3.45 | 0.93 | 1.06 | 0.0 | 0.61 |
| Rater   | 0.15 | 0.57 | 3.26 | 0.93 | - | - | - |

All items were in valid condition shown by the inexistence of outlier or mis-fit items. An item was classified as fit, if the logit score exceeds +2SD and -2SD, while misfit items were determined by outfit score MNSQ, ZSTD, and Pt Mea Corr [21]. The average score of MNSQ outfit obtained from
MFRM items in this research approaching 1 (1.06) showed that there was no misfit item with the model supported with the average of ZSTD score was 0.0 and the average of Point Measure (Pt. Mea) was 0.61. It was shown that all items were appropriate with latent variable that had been agreed.

According to the separation score of ratee (2.50), teacher participants was classified into three groups, they are the good, moderate, and poor teachers. This classification was shown in Figure 1. There are three teachers (15%) who are categorized as good teachers, namely T19, T18 and T3. The majority of teachers are in moderate ability (13 teachers or 65%). While the group of poor teachers consists of 4 teachers.

![Figure 1. Variable map of ratee and item](image)

The MFRM result analysis involving three facets that were 20 teachers’ PCK (rate), PCK of rubric items for specific topic (item) and 6 raters was shown on Picture 1. The first column showed logit scale showing the measurement unit to all three facets. The second column, third, and fourth consequently were the logit measurement distribution to ratees, items, and raters. The focus discussion in this article just on the first to third column where logit of ratee and items are shown. All the distributions were modeled with a mean of zero logit (0.0) that meant the average of teachers’ PCK competence, the average of difficulty level item and an average of severity rater [21]. All logit distributions to all three facets were ordered from the lowest to the highest.

3.3. Teachers’ PCK on each aspect
The MFRM analysis to every aspect of PCK shown on the third column of Picture 1 described the item difficulty level. The distribution of measurement logit was ordered from the lowest at the bottom position and vice versa. Based on the SD score (0.96), the item was classified into three groups, they are the hard, moderate, and easy items. The result of the classification was shown on Table 4.

Based on Table 4, the difficult item was the PCK item on the representations and conceptual teaching strategies aspects. The difficult item of MFRM analysis indicated that almost all teachers had problems (did not master) the particular item, because only teachers with high ability could perform or master the particular item. In other words, the competence of the representation and conceptual teaching strategies of all teachers was still low, especially on items with codes S06, S07, S09, and S10. The four items were the items showing the use of computer animation and role playing demonstration in learning the matter characteristic to understand the sub-microscopic matter. All the items in curricular saliency aspect and learner’s prior knowledge were classified in items with moderate difficulty level. Meanwhile, the easiest item, P04, was the item to examine the learners’ prior knowledge. Since item P04 was the easiest item, indicates that the teacher
already understands that students have understood in relation to the nature of matter, especially in macroscopic nature, such as solids are rigid and liquid forms follow the container. Meanwhile, items S06; S07; S09; and S10 were the most difficult items, it indicated that they did not understand the learning strategies that could overcome the problem, namely submikroskopis nature related with particle arrangement in solid, liquid and gas. Thus, the majority of teachers then employed the learning strategy with item codes S8 and S11, learning strategy that employed real objects to classify matter into solid, liquid, and gas based on the apparent physical characteristics (macroscopic nature) [22]. Moreover, the learning process employing real objects focused on the macroscopic nature was more appropriate for learning Science in elementary school level, not in middle school because middle school students should already been introduced to something abstract [23].

| No | PCK Aspects                          | Difficult item         | Moderate item | Easy item   |
|----|--------------------------------------|------------------------|---------------|-------------|
| 1  | Curricular saliency                  | C01; C02 & C03         |               |             |
| 2  | Learner prior knowledge              | P04                    |               |             |
| 3  | What makes a topic easy or difficult to understand | D05                    |               |             |
| 4  | Representations & conceptual teaching strategies | S06; S07; S09 & S10   | S08; S11; S12 |             |

The result of the PCK competence mapping in this research had illustrated clearly the weak or strong PCK ability of the teacher participants considering the low competence of Indonesian teachers [24], especially science teachers. Therefore, the result of MFRM analysis to the teachers’ PCK could be a reference for science teacher educators in designing the teacher training program. Although this research had the limitation of only investigating the teachers’ PCK on only one topic, the result of this research had big implication to the research related with the teacher certification program in Indonesia that aimed to develop teachers’ professionalism.

4. Conclusion
MFRM analysis presented clear information of PCK ability of 20 middle school science teacher participants in this research. The PCK abilities could be classified into three, they were: good, moderate, and poor teachers’ PCK. In conclusions, in-depth analysis to teachers’ PCK aspects with MFRM showed that generally teachers involved as respondents possessed low PCK ability in representation and conceptual teaching strategies. Thus, this aspect still needs to be improved in the teacher training and education program in Indonesia.

Acknowledgement
The authors thanks to teacher participants in Yogyakarta spesial region and all raters who are lecturers in the science education study program of Mathematics and natural science faculty, yogyakarta state University. This research is one of the researches in the doctoral dissertation research funded by the Director General of Research and Development Strengthening the Ministry of Research, Technology and Higher Education of the Republic of Indonesia No. 3/E/KP/2018.

References
[1] Sumintono, B. & Subekti, N. 2014 teacher In-Service Training and Retraining in Indonesia. Chapter of International Handbook of Teacher Education, Training and Re-training Systems in Modern World, editors: Konstantinos G. Karras Charl C. Wollhuter, ISBN 978-9963-2093-4-7, Studies and Publishing, Nicosia, Cyprus, 2014
[2] UURI. (2005). Undang-Undang Republik Indonesia Nomor 14 Tahun 2005 tentang Guru dan Dosen [Law of the Republic of Indonesia No. 14 Year 2005 on Teachers and Lecturers]. Jakarta: BP. Dharma Bhakti

[3] Shulman, L.S. (1986). Those who understand: Knowledge growth in teaching. Educational Researcher, 15, 4-14.

[4] Shulman, L.S. (1987). Knowledge and teaching: foundations of the new reform. Harvard Educational Review.

[5] National Research Council. 1996. National Science Education Standard. Washington, DC: National Academic Press.

[6] Mishra, P. & Koehler, M.J., (2006), Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge, Teachers College Record 108 (6), 1017-1054.

[7] Van Driel & Berry, 2012, Teacher Professional Development Focusing on Pedagogical Content Knowledge. Educational Researcher, Vol 41, No.1, pp26-28.

[8] Kuhn T 1998 Density matrix theory of coherent ultrafast dynamics Theory of Transport Properties of Semiconductor Nanostructures (Electronic Materials vol 4) ed E Schöll (London: Chapman and Hall) chapter 6 pp 173–214

[9] Van Driel, J.H., Verloop, N., & De Vos, W. (1998). Developing science teachers’ pedagogical content knowledge. Journal of Research in Science Teaching, 35(6), 673–695.

[10] Veal, W., & MaKinster, J. (1999). Pedagogical Content Knowledge Taxonomies. Electronic Journal of Science Education, 3(4) Article Two.

[11] Gess-Newsome, J. (2015), A Model of Teacher Professional knowledge and skill including, Result of the thinking from the PCK summit. Chapter of e-book Re-examining Pedagogical Content Knowledge in Science Education, Edited by Amanda Berry, Patricia Friedrichsen, John Loughran, Routledge, New York.

[12] Park, S., & Chen, Y. (2012). Mapping out the integration of the components of pedagogical content knowledge (PCK): Examples from high school biology classrooms. Journal of Research in Science Teaching, 49(7), 922–941

[13] Snir, J., Smith, C. L. & Raz, G. (2003). Linking phenomena with competing underlying models: A software tool for introducing students to the particulate model of matter. Science Education, 86, 694-830

[14] Boz, N. & Boz, Y. 2008 A qualitative case study of prospective chemistry teachers’ knowledge about instructional strategies: Introducing particulate theory. Journal of Science Teacher Education, 19(2), 135-156.

[15] Tsai, C.C. (1999). Overcoming junior high school students’ misconceptions about microscopic views of phase change: A study of an analogy activity. Journal of Science Education and Technology, 8(1), 83-91.

[16] Baxter, J. A., & Lederman, N. G. (1999). Assessment and measurement of pedagogical content knowledge. In J. Gess-Newsome and N. G. Lederman (Eds.), Examining pedagogical content knowledge (pp. 147-163). Boston, MA: Kluwer

[17] Merriam, S. B. (1998). Qualitative research and case study applications in education. San Francisco: Jossey-Bass.

[18] Boone, W. J., Stauber, J.R. and Yale, M.S. (2014). Rasch Analysis in the Human Sciences.Dordrecht: Springer.

[19] Engelhard Jr, G. (2013). Invariant measurement: Using Rasch models in the social, behavioral and health sciences. New York: Routledge

[20] Bond,T.G.and Fox,C.M.(2015). Applying The Rasch Model, fundamental measurementin the human sciences 3rd edition. NewYork:Routledge.

[21] Perera, Sumintono & Na, (2018), The Psychometric Validation of The Principal Practices Questionaire Based on Item Response Theory, International Online Journal of Educational Leadership, 2018 Vol. 2, No. 1, 21-38

[22] Boz, N. & Boz, Y. 2008 A qualitative case study of prospective chemistry teachers’ knowledge about instructional strategies: Introducing particulate theory. Journal of Science Teacher Education, 19(2), 135-156.

[23] Toplis, R.. 2010. Learning to Teach Science in the Secondary School: A Companion to School Experience (Learning to Teach Subjects in the Secondary School Series), 3rd Edition, by Routledge, published in the Taylor & Francis e-Library
[24] De Ree, J., Muralidharan, K., Pradhan, M. & Rogers, H., 2018. Double for Nothing? Experimental Evidence on an Unconditional Teacher Salary Increase in Indonesia. *The Quarterly Journal of Economics*, Volume 133, Issue 2, 1 May 2018, 993–1039.