Developing Competency Evaluation of Pre-service Science Teachers in Industrial Revolution 4.0: Revealing Pedagogic and Professional Competencies

by Lukman Abdul Rauf Laliyo

Research Article https://doi.org/10.12973/ijem.8.2.347

International Journal of Educational Methodology
Volume 8, Issue 2, 347 - 362.

ISSN: 2469-9632
https://www.ijem.com/

Developing Competency Evaluation of Pre-service Science Teachers in Industrial Revolution 4.0: Revealing Pedagogic and Professional Competencies

Endang Susantini* Universitas Negeri Surabaya, INDONESIA

Badrun Kartowagiran Universitas Negeri Yogyakarta, INDONESIA

Syukrul Hamdi Universitas Negeri Yogyakarta, INDONESIA

Samsul Hadi Universitas Negeri Yogyakarta, INDONESIA

Amat Jaedun Universitas Negeri Yogyakarta, INDONESIA

I Gede Astra Wesnawa Universitas Pendidikan Ganesha, INDONESIA

Dadang Sunendar Universitas Pendidikan Indonesia, INDONESIA

Lukman Abdul Rauf Laliyo Universitas Negeri Gorontalo, INDONESIA

Received: January 25, 2022 ▪ Revised: April 12, 2022 ▪ Accepted: May 11, 2022

Abstract: This study aimed to develop an instrument to assess science teacher candidates' professional and pedagogical competence in the Industrial revolution 4.0. The instrument consisted of 30 questions and was used in the main study, which was analyzed using the Rasch model to unravel the reliability and empirical validity. The questions were developed based on a predetermined grid including the skills and literacy aspects used in the industrial revolution 4.0. The 30 revised questions were then tested on 60 students of Science Education at two State universities in Indonesia who were in their third year of study times. The Rasch model was employed to test its reliability and empirical validity that included Wright map analysis, item difficulty level, distractor analysis, and item suitability with the Rasch model. The results showed that the reliability value of the professional aspect was .95 and classified as an excellent category. Moreover, the pedagogic aspect obtained .93 and was classified as a very good category. Only one question was the most difficult and did not fit the Rasch model, while the others already fitted. This reliable and valid instrument is suggested to be useful in assessing pre-service science teachers' competence.

Keywords: *Industrial revolution, Rasch model, teacher competence, test development.*

To cite this article: Susantini, E., Kartowagiran, B., Hamdi, S., Hadi, S., Jaedun, A., Wesnawa, I. G. A., Sunendar, D., & Laliyo, L. A. R. (2022). Developing competency evaluation of pre-service science teachers in industrial revolution 4.0: Revealing pedagogic and professional competencies. *International Journal of Educational Methodology*, *8*(2), 347-362. https://doi.org/10.12973/ijem.8.2.347

Introduction

In the 21st century, technology and information have been increasing in all aspects of life. In the 1760s, the first industrial revolution was sparked and indicated the development of tools used for industrial activities. In the 2000s, it reached the fourth industrial revolution or known as the industrial revolution 4.0. When compared to the previous era, the industrial revolution 4.0 is growing exponentially. In addition, such development occurs throughout the industry in almost every country. This makes the scope of the industrial revolution 4.0 wider to include transformations in all systems of production, management, and government (Xu et al., 2018). Exponential technological developments in industrial revolution 4.0 include artificial intelligence (AI), biotechnology, nanomaterials, environmentally friendly natural resources, and genetic technology (Rahiem, 2020).

There are four principles characterizing the industrial revolution 4.0, namely systemic impact, empowerment, future orientation, and usefulness (Philbeck, 2017). These principles are broadly translated into skills that must be possessed by every individual in the 4.0 industrial revolution era, including complex problem solving, critical thinking, creativity, people management, coordinating with others, emotional intelligence, judgment and decision making, service orientation, negotiation, and cognitive flexibility (World Economic Forum, 2016).

Educational institutions have an important role in the formation of human resources that can compete in the industrial revolution 4.0 (Alda et al., 2020). Maryanti et al. (2020) stated that educational institutions must be able to graduate

* **Corresponding author:**
Endang Susantini, Universitas Negeri Surabaya, Jl. Lidah Wetan, Surabaya, Indonesia. ✉ endangsusantini@unesa.ac.id

individuals who are ready to face the challenges so that they become more productive individuals. This is inseparable from the role of the learning process that occurs in formal and non-formal education. Teachers as facilitators as well as tutors in the classroom are required to have the skills required in the industrial revolution 4.0, e.g., ably providing appropriate provisions for their students in facing the development of the industrial revolution 4.0 (Indira et al., 2020). Moreover, they are suggested to be able to develop basic competencies, including professional, pedagogic, social, and personality competencies, which are relevant to the industry revolution 4.0 skills.

Industrial revolution 4.0 demands the use of digital technology in the learning process or known as a cyber system (Ellitan, 2020). This system is able to make the learning process take place regularly without space and time limits. Indonesia can be said to be slower in responding to the industrial revolution 4.0 than the neighboring countries such as Malaysia and Singapore. The new 4.0 education system reverberates this year. Therefore, the government must provide adequate facilities in welcoming the era of Education 4.0. As the front line in the world of education, teachers must upgrade their competencies in the face of the 4.0 Education era. Students respond to any critical phenomena faster than teachers. In the end, school output can produce quality human beings commensurate with industrial revolution 4.0.

As a place to produce qualified teacher candidates, the university should begin its innovative assessment of the pre-service teachers' pedagogic and professional competence (Akimova et al., 2020). The two competencies are indeed necessary for keeping up the quality of education 4.0. Pedagogic competence refers to any knowledge and practices of the teaching and learning process including but not limited to instructional design, curriculum development, producing authentic lesson plans, and vice versa (Bakar et al., 2020; Juhji & Nuangchalerm, 2020). These competencies were obligatory to be mastered by the pre-service teachers nowadays since the pre-requisite to deal with the current job vacancy has been so complex, e.g., the need for critical thinking and computerization skills (Abdullah et al., 2020). Moreover, professionalism should be permeated completely to hinder them from job burnout and absenteeism. Professional aspects might give various perspectives on how to manage a classroom, school programs, and student problems (Macaro et al., 2020). Therefore, these two competencies should be appropriated by the pre-service teachers at the very beginning.

By looking at the above seriousness, the Indonesian government is trying to deal with the challenges by mounting the competence of pre-service teachers. One of the methods used to improve teacher competence is conducting a preliminary study to determine the competence of pre-service teachers by providing appropriate evaluations. The evaluation of the pre-service teachers covers all subjects, however, the present study focused on science subjects. Therefore, in this study, the development of an instrument for evaluating the competence of pre-service science teachers, especially in the professional and pedagogical competencies, was carried out in coping with the industrial revolution 4.0. The present study aimed to reveal the reliability and empirical validity of the developed instrument for pre-service science teachers' competence evaluation.

Literature Review

Teacher Competencies in the Industrial Revolution 4.0

The Industrial Revolution was first introduced in the 1780s with the development of the steam engine that made the industrial world more productive (Xu et al., 2018). In the 1800s, there was a second industrial revolution indicated by the development of electrical energy and mass production. The third industrial revolution occurred around the 1900s with the development of information and communication technology devices and the term globalization began to emerge. The development of information and communication technology in the era of the industrial revolution 3.0 made production more efficient. In the early 2000s, the industrial revolution 4.0 was introduced and there was a combination of several technologies and knowledge. Klaus Schwab at the World Economic Forum 2016 stated that the industrial revolution 4.0 was a new technology that became a combination of the real, digital, and biological worlds, which had an impact on all disciplines, economic aspects, and industrial fields (Kopotun et al., 2020). He added that even in the era of the industrial revolution 4.0, it was also specified by ideas that challenged humans.

In industrial revolution 4.0, there is a massive and exponential increase in the use of technology. Digitization occurs in all aspects of life and so does computerization (Gleason, 2018). Industrial revolution 4.0 involves technological and multidisciplinary convergence that is different from the past. This difference has a multi-systemic impact where new technologies will give rise to thoughts regarding how the technology should be used and its benefits for human welfare (Philbeck, 2017).

The industrial revolution 4.0 process will control and shape all social groups including young people, adults, and workers around the world. One of the impacts is that there is no limitation of space and time to work. In the field of education, the industrial revolution 4.0 also conveys a significant impact. Educational institutions must be able to equip their students with the skills required by the industrial revolution 4.0 such as interdisciplinary creativity, critical thinking, and problem-solving skills. Pre-service teachers who will later become educators also need to have these competencies. The competencies that teachers need to have in the era of the industrial revolution 4.0 include (Wahyuni, 2018) educational competence, technology commercialization competence, globalization competence, competence to analyze future strategies, and counselor competence.

Evaluation Instrument Development

Evaluation is a process carried out to find out and measure something using certain measuring instruments (Muryadi, 2017). Something measured can be in the form of whether cognitive, psychomotor, or social skills. Valid measurements are used to assess what you want to assess and have been standardized as evidenced by the appropriate validity and reliability test (Arifin, 2009). The measuring instrument can be a test or a non-test instrument. Measuring students' abilities can use a specially designed test according to the grid that has been determined.

One of the test instruments commonly used is the multiple-choice test. Among many types of tests, multiple-choice tests remain a good option to measure cognitive skills and knowledge, and their use is effective and efficient (Haladyna et al., 2019). Multiple-choice tests can be used for a large number of respondents and the items developed are also flexible (Türkoguz, 2020). The multiple-choice test consists of one question with several answer options, among the answer options there is the most correct answer and the others are distractors. The process of developing multiple-choice tests can be carried out in the following steps: 1) planning; 2) carrying out trials; 3) conducting validation of test instruments; 4) measuring the reliability of the test; and, 5) undertaking data analysis (Wartoni & Benyamin, 2020). The planning stage consists of determining the purpose of the test, determining the form of the test, developing a test grid, compiling test items, and making revisions.

Multiple-choice tests that have been developed must be analyzed for quality. The developed test is expected to produce objective and accurate conclusions. Tests that are not of good quality will result in misinterpretations and might affect subsequent policies (Arifin, 2009). Analysis of the test was carried out as an aggregate and as for each item. The most basic test analysis includes validity and reliability. A validity test is measuring whether the developed test can measure what should be measured or not and a reliability test is the consistency of the tests carried out by respondents (Fraenkel et al., 2012). Validity includes content validity that covers instrument format, criteria validity, and constructs validity.

Item analysis can be conducted using the classical method or Rasch modeling. Analysis of items based on classical theory consists of reliability tests using the Cronbach-Alpha test, discriminating power of questions, level of difficulty, and distracting analysis using certain formulas (Arifin, 2009). The item analysis with Rasch modeling uses dichotomous data that can interpret the items and respondents' abilities in detail (Nuryanti et al., 2018). The dichotomy data are processed using special software called Winsteps. The data obtained from the software is in the form of logarithmic values presented in the form of tables and scales. The analysis of the Rasch model using Winsteps application includes an item measure that shows the level of difficulty of each item and its reliability; misfit items that indicate whether the items fit the Rasch model; Wright's map showing the relationship between each item and the respondents; and Pearson measure that shows the ability of each respondent.

Methodology

This study was conducted using two phases namely pre-study and main study. The pre-study covered generating an initial draft of the instrument, while the main study focused on revealing the reliability and empirical validity test. In coping with the pre-study phase, there were 90 questions developed of each pedagogic and professional competency. The 90 questions were drawn based on the predetermined grids namely science content which includes the form of matter and the form of objects, solutions, acids and bases, energy, organization of living things, food chains, environmental changes, biogeochemical cycles, the motion of revolution and rotation of the earth and moon, structure and function of plant tissue, blood tissue, simple planes, digestive system and additives, respiratory system, locomotion system, excretory system, biotechnology, reproductive system, development and growth, cell reproduction, genetic substance and heredity, and adaptation of creatures life. All the grids were integrated into ten components of the industrial revolution 4.0 namely critical thinking, cognitive ability, complex problem solving, creativity, judgment and decision making, people management, coordinating with others, emotional intelligence, service orientation, and negotiation. The questions were in the form of multiple-choice with five answer choices. The initial instrument draft was then assessed by science experts to look at the composition of the questions. The experts argued that the number of the questions was too many so they suggested being reduced to 30 questions for each pedagogic and professional competence.

The instrument was then revised to cope with the experts' suggestions. The revised draft included 30 questions representing ten components of industrial revolution 4.0 for each competence of pedagogy and professionalism. Table 1 shows the blueprint of the revised draft.

Table 1. Blueprint of the Revised Instrument

Competence	Components of Industrial Revolution 4.0	Item Number
Pedagogy	Complex problem-solving	1-3
	Critical thinking	4-6
	Creativity	7-9
	People management	10-12
	Coordinating with others	13-15
	Emotional intelligence	16-18
	Judgment and decision making	19-21
	Service orientation	22-24
	Negotiation	25-27
	Cognitive flexibility	28-30
Professionalism	Complex problem-solving	31-33
	Critical thinking	34-36
	Creativity	37-39
	People management	40-42
	Coordinating with others	43-45
	Emotional intelligence	46-47
	Judgment and decision making	48-51
	Service orientation	52-54
	Negotiation	55-57
	Cognitive flexibility	58-60

There were 60 questions (30 questions for pedagogy competence and 30 questions for professional competence) that represent ten components of industrial revolution 4.0 in science tested on 60 students of Science Education at two State universities in Indonesia. The 60 respondents had fulfilled the criteria of the statistical test (Pallant, 2020) and, therefore, the study could be undertaken. Moreover, the respondents were limited to several research participant criteria namely pursuing the sixth semester of the study times and having been enrolled in Innovative Learning, Learning Theory, Learning Media Development, and Assessment Courses. The respondents had no scores below the average in those four initial courses. The four courses were necessary for providing the contents of the evaluation sheet that covered magnetic and electricity, solutions, school science analysis, and environments. Respondents' answers were empirically validated, the reliability value of the questions was calculated, and each item was analyzed using the Rasch Model using the Winsteps application. The item analysis included Wright's map (Logic scale), item measure order, item misfit order, and distractor analysis.

The resulted scores of the reliability test could be categorized using the following description (see Table 2).

Table 2. Categorization of Reliability Scores

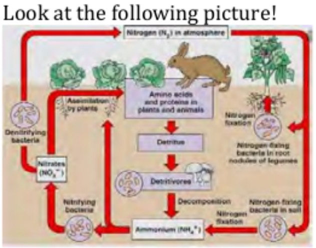
Scores	Category
< .67	Low
.67 – .80	Fair
.81 – .90	Good
.91 – .94	Very Good
>.94	Excellent

Wright's map analysis showed two things, namely the difficulty of the questions and the ability of the respondents. The higher the position of the item, the higher the difficulty level, the lower the position of the item, the easier the question. Likewise, the higher the position of the respondent, the higher the ability or the higher the score obtained. Item measure analysis was used to see the difficulty of the questions. The classification of the difficulty level of the question could be seen from the comparison of the measured value with the Standard Deviation (SD).

Findings / Results

The study suggested that 30 questions developed had been declared valid based on theoretical validation by three science education experts. The following were examples of professional and pedagogic questions that were tested for pre-service science teachers.

Look at the following picture!



Based on the picture above, the right way to maintain the sustainability of the nitrogen cycle is...

A. maintaining animal population as high as possible to supply organic nitrogen
B. maintaining plant population as high as possible to supply inorganic nitrogen
C. maintaining a high population of soil microorganisms that play a role in nitrification
D. maintaining a high population of soil microorganisms that play a role in denitrification
E. maintaining the balance of the population of nitrifying and denitrifying bacteria

Figure 1. One of the Questions about the Professional Competence

Students in class A quickly get bored when the teacher explains the material in class. Not a few of them ask permission to go to the bathroom and return to class after a long time and some are truant. Mrs. Della is a science teacher designed a study on pollution material that aims to enable students to analyze the occurrence of pollution and its impact on the environment. Knowing the character of the students in class A, the appropriate learning activity for Mrs. Della is…

A. providing projects related to pollution in the surrounding environment
B. observing pictures of environmental pollution
C. watching video shows about the pollution that occurs
D. conducting discussions and ask questions to reduce the impact of pollution
E. providing pollution articles for analysis and resume

Figure 2. One of the Questions about the Pedagogic Competence Professional Competence

In connection with Table 3, the reliability of the respondents and the items showed the reliability value of the items of .95, which was classified as excellent. This meant that the questions developed were very consistent in measuring students' abilities. Moreover, it was also known that the value of the separation item was 4.48, which was classified as very good, meaning that students were varied enough to be able to detect the items given.

Table 3. Summary of Item Measure

	Total Score	Count	Measure	Model Error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	31.5	60.0	.00	.37	.99	.1	1.03	.1
SD	18.4	.0	1.82	.11	.09	.6	.40	1.0
Max.	57.0	60.0	3.73	.72	1.16	1.9	2.38	2.2
Min.	2.0	60.0	-2.97	.27	.77	-.8	.33	-1.5
Real RMSE .40		True SD 1.78		Separation 4.48		Item Reliability		.95
Model RMSE .39		True SD 1.78		Separation 4.48		Item Reliability		.95
S.E. of Item Mean = .34								

Figure 3 shows an image of Wright's map that compared respondents with items. *Person* was the respondent who was working on the question, and the item was the question item. The higher the position of the item, the higher the level of difficulty, which was question 18. Meanwhile, the lower the position of the item, the easier the question, namely questions 11 and 21. The measured value of each item could be observed in the item measure table (see Figure 3).

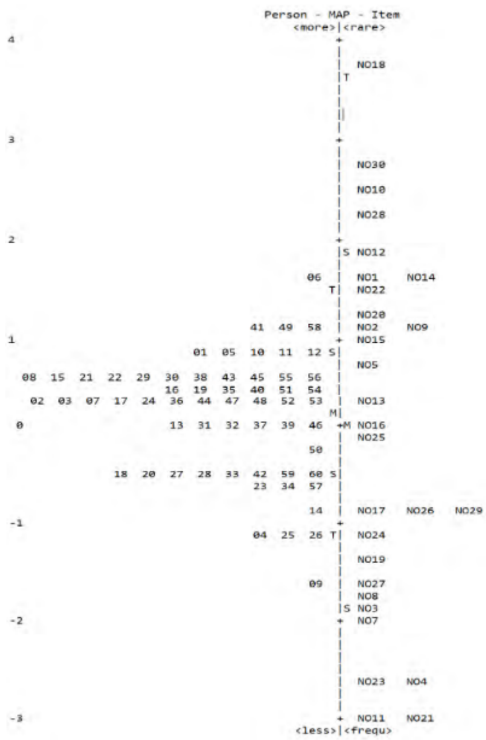


Figure 3. Wright's Map Points to Professional Competence

In the following Table 4, it can be seen the level of difficulty of the questions. The most difficult question was number 18 with a measured value of 3.73 and could only be answered correctly by 2 out of 60 respondents. While the easiest questions were questions 11 and 21, which could be answered correctly by 57 out of 60 people.

Table 4. Items of Professional Competence

Entry Number	Total Score	Total Count	Measure	Model Se	Infit Mnsq	Infit Zstd	Outfit Mnsq	Outfit Zstd	Pt-Measure Corr.	Pt-Measure Exp	Exact Match Obs%	Exact Match Exp%	Item
18	2	60	3.73	.72	1.02	.3	1.02	.3	.05	.10	96.7	96.7	18
30	5	60	2.74	.47	1.08	.3	2.38	2.2	-.15	.15	91.7	91.6	30
10	6	60	2.53	.44	1.15	.6	2.05	2.0	-.24	.17	90.0	90.0	10
28	8	60	2.20	.39	1.03	.2	1.62	1.6	.06	.19	86.7	86.6	28
12	10	60	1.92	.35	.93	-.2	.83	-.5	.33	.21	83.3	83.3	12
1	13	60	1.58	.32	.98	.0	.87	-.5	.29	.23	76.7	78.4	1
14	13	60	1.58	.32	1.10	.6	1.29	1.2	.04	.23	76.7	78.4	14
22	14	60	1.48	.31	.97	-.1	1.21	.9	.23	.24	78.3	76.8	22
20	16	60	1.29	.30	.99	.0	1.01	.1	.26	.25	71.7	73.6	20
2	18	60	1.11	.29	1.05	.4	1.06	.4	.19	.26	66.7	70.5	2
9	18	60	1.11	.29	1.09	.8	1.12	.7	.12	.26	73.3	70.5	9
15	19	60	1.03	.29	.91	-.8	.85	-1.0	.41	.27	68.3	69.1	15
5	23	60	.71	.28	1.07	.9	1.09	.7	.17	.28	63.3	64.3	5
6	27	60	.41	.27	.95	-.7	.95	-.5	.37	.30	71.7	62.5	6
13	29	60	.26	.27	1.08	1.0	1.07	.8	.19	.30	55.0	62.2	13
16	33	60	-.03	.27	1.16	1.9	1.15	1.6	.08	.31	53.3	63.8	16
25	35	60	-.18	.28	1.10	1.1	1.14	1.3	.15	.31	63.3	65.2	25
17	43	60	-.83	.30	.96	-.3	.87	-.7	.38	.29	70.0	73.1	17

Table 4. Continued

Entry Number	Total Score	Total Count	Measure	Model Se	Infit		Outfit		Pt-Measure		Exact Match		Item
					Mnsq	Zstd	Mnsq	Zstd	Corr.	Exp	Obs%	Exp%	
26	44	60	-.92	.31	.88	-.8	.83	-.9	.47	.29	76.7	74.4	26
29	44	60	-.92	.31	.92	-.5	.83	-.9	.43	.29	73.3	74.4	29
24	46	60	-1.11	.32	.93	-.4	.92	-.3	.38	.28	76.7	77.1	24
19	48	60	-1.33	.34	.98	.0	.92	-.2	.32	.27	78.3	80.2	19
27	50	60	-1.57	.36	.99	.0	.87	-.4	.30	.26	85.0	83.3	27
8	51	60	-1.70	.37	.90	-.3	.75	-.7	.42	.25	85.0	84.9	8
3	52	60	-1.85	.39	1.01	.1	1.03	.2	.22	.24	86.7	86.6	3
7	53	60	-2.01	.41	.83	-.5	.63	-1.0	.49	.23	88.3	88.3	7
4	56	60	-2.65	.53	.77	-.4	.33	1.5	.59	.18	93.3	93.3	4
23	56	60	-2.65	.53	.96	.1	.97	.1	.21	.18	93.3	93.3	23
11	57	60	-2.97	.60	.99	.2	.86	.0	.19	.16	95.0	95.0	11
21	57	60	-2.97	.60	.86	.1	.42	.9	.44	.16	95.0	95.0	21
MEAN	31.5	60.0	.00	.37	.99	.1	1.03	.1			78.8	79.4	
SD	18.4	.0	1.82	.11	.09	.6	.40	1.0			11.7	10.518	

Each item was also analyzed for its suitability with the Rasch model, namely by looking at the table of misfit items (see Table 5). An item was said to fit the Rasch model if: The outfit mean square (MNSQ) value was less than 1.5 and more than .5; standard outfit value-Z (ZSTD) -2.0 < ZSTD < +2.0; and the value of point measure correlation (Pt Measure Cor.) .4 < Pt Measure Cor. < .85. If one item was found where the MNSQ and Pt Measure Cor. values did not meet the criteria but the ZSTD value met the criteria, the item was still considered fit, meaning that the item was maintained.

In accordance with Figure 5, question 30 was not fit because the MNSQ value was more than 1.5 and the ZSTD value was more than 2. Question 10 was fit because the ZSTD value was equal to 2 even though the MNSQ value was more than 1.5. Question 28 fitted because the ZSTD value was less than 2 even though the MNSQ value was more than 1.5. Questions 10 and 28 could be corrected to get an MNSQ score that was less than 1.5. In addition, it was also known that the point-measure correlation (PT-MEASURE CORR) value showed the differentiating power of the items. It could be seen that the PT-MEASURE CORR score was mostly still below .4 and there were only 7 items that got a score above .4. This showed that the discriminatory power of the questions was not good. Items could be maintained if the ZSTD score met the requirements, even though the MNSQ and PT-MEASURE CORR scores did not meet the requirements. Overall, of the 30 items, there were 7 questions that met the three fit item criteria and were said to be invalid. The question could be replaced or not included in the next test. There were 20 questions that met the MNSQ and ZSTD criteria but did not meet the PT-MEASURE CORR criteria so the 20 questions could be defended.

Table 5. Misfit Items of Professional Competence

Entry Number	Total Score	Total Count	Measure	Model Se	Infit		Outfit		Pt-Measure		Exact Match		Item
					Mnsq	Zstd	Mnsq	Zstd	Corr.	Exp	Obs%	Exp%	
30	5	60	2.74	.47	1.08	.3	2.38	2.2	A-.15	.15	91.7	91.6	30
10	6	60	2.53	.44	1.05	.6	2.5	2.0	B-.24	.15	90.0	90.0	10
28	8	60	2.20	.39	1.03	.2	1.62	1.6	C .06	.19	86.7	86.6	28
14	13	60	1.58	.32	1.10	.6	1.29	1.2	D .04	.23	86.7	78.4	14
22	14	60	1.48	.31	.97	-.1	1.21	.9	E .23	.24	83.3	76.8	22
16	33	60	-.03	.27	1.16	1.9	1.15	1.6	F .08	.31	76.7	63.8	16
25	35	60	-.18	.28	1.10	1.1	1.14	1.3	G .15	.31	76.7	65.2	25
9	18	60	1.11	.29	1.09	.8	1.12	.7	H .12	.26	78.3	70.5	9
5	23	60	.71	.28	1.07	.9	1.09	.7	I .17	.26	71.7	83.3	5
13	29	60	.26	.27	1.08	1.0	1.07	.8	J .19	.30	66.7	78.4	13
2	18	60	1.11	.29	1.05	.4	1.06	.4	K .19	.26	73.3	78.4	2
3	52	60	-1.85	.39	1.01	.1	1.03	.2	L .22	.24	68.3	76.8	3
18	2	60	3.73	.72	1.02	.3	1.02	.3	M .05	.10	63.3	73.6	18
20	16	60	1.29	.30	.99	.0	1.01	.1	N .26	.25	71.7	70.5	20
11	57	60	-2.97	.60	.99	.2	.86	.0	O .19	.16	55.0	70.5	11
27	50	60	-1.57	.36	.99	.0	.87	-.4	O .30	.26	53.3	69.1	27
1	13	60	1.58	.32	.98	.0	.87	-.5	N .29	.23	63.3	64.3	1
19	48	60	-1.33	.34	.96	.0	.92	-.2	M .32	.27	70.0	62.5	19
23	56	60	-2.65	.53	.96	.1	.97	.1	L .21	.18	76.7	76.9	23
17	43	60	-.83	.30	.95	-.3	.87	-.7	K .38	.29	73.3	74.3	17

Table 5. Continued

Entry Number	Total Score	Total Count	Measure	Model Se	Infit Mnsq	Infit Zstd	Outfit Mnsq	Outfit Zstd	Pt-Measure Corr.	Pt-Measure Exp	Exact Match Obs%	Exact Match Exp%	Item
6	27	60	.41	.27	.93	-.7	.95	-.5	J .37	.30	76.7	76.6	6
24	46	60	-1.11	.32	.93	-.4	.92	-.3	I .38	.28	78.3	78.9	24
12	10	60	1.92	.35	.92	-.2	.91	-.5	H .33	.21	85.0	84.6	12
29	44	60	1.03	.31	.91	-.5	.90	-.9	G .43	.29	85.0	85.0	29
15	19	60	-1.70	.29	.88	-.8	.88	-1.0	F .41	.27	68.3	68.3	15
8	51	60	-.92	.37	.86	-.3	.86	-.7	E .42	.25	88.3	84.9	8
26	44	60	-2.97	.31	.86	-.8	.85	-.9	D .47	.29	93.3	74.4	26
21	57	60	-2.97	.60	.84	-.1	.75	-.9	C .44	.16	93.3	95.0	21
7	53	60	-2.01	.41	.83	-.5	-83	-1.0	B .49	.23	88.3	88.3	7
4	56	60	-2.65	.53	.77	-.4	.33	-1.5	A .59	.18	93.3	93.3	4
MEAN	31.5	60.0	.00	.37	.99	.1	1.03	.1			78.8	79.4	
SD	18.4	.0	1.82	.11	.09	.6	.40	1.0			11.7	10.5	

In developing questions, one thing to be considered was the distractor in the answer options. The distractor analysis could be seen from the average ability value (see Table 6). If the trend was up for each answer option, the distractor worked well. Questions 18, 30, and 10 showed an inconsistent order of average ability values, which meant the distractor option did not work well. In Table 4, it could be seen that option E in question 30 was at the top with an average ability value of -.13, and option D was at the bottom with an average ability value of -.14, so the distractor on this question was not good. The minus value on average ability indicated that there were respondents who had low abilities who could answer the correct answer so it could be said that the respondent only guessed correctly. Questions whose distractors worked well might cover questions 12 and 1. In question 12, it was known that option E was in first place with an average ability value of -.03, and option B was the correct answer at the bottom with an average ability value of .67 so the trend of the average ability value increased, then the distractor worked well.

Table 6. Distribution of Distractors for Each Item

Entry Number	Data Code	Score Value	Data Count	Data %	Average Ability	S.E. Mean	Outfit MNSQ	PTMEA Corr.	Item
18	C	0	14	23	-.08	.16	.8	-.21	18
	A	0	19	32	.16	.15	1.0	-.02	
	E	0	13	22	.20	.21	1.1	.02	
	B	0	12	20	.45	.21	1.3	.21	
	D	1	2	3	.35*	.11	1.0	.05	
30	E	0	12	20	-.13	.23	.9	-.23	30
	A	0	10	17	.01	.19	.9	-.12	
	D	0	10	17	.17	.18	1.0	-.01	
	C	0	23	38	.49	.10	1.3	.37	
	B	1	5	8	-.14*	.46	2.5	-.15	
10	B	0	2	3	-.71	.94	.5	-.25	10
	A	0	4	7	-.22	.39	.7	-.16	
	C	0	9	15	.00	.25	.9	-.11	
	E	0	39	65	.38	.09	1.2	.41	
	D	1	6	10	-.30	.23	2.2	-.24	
28	E	0	2	3	-.33	.34	.5	-.14	28
	D	0	50	83	.18	.09	1.0	.01	
	C	1	8	13	.28	.34	1.7	.06	
12	E	0	7	12	-.03	.26	.9	-.11	12
	A	0	42	70	.10	.10	1.0	-.19	
	C	0	1	2	.23		.9	.01	
	B	1	10	17	.67	.17	.8	.33	
1	E	0	6	10	-.12	.50	1.2	-.15	1
	B	0	21	35	-.07	.14	.8	-.28	
	C	0	12	20	.29	.17	1.1	.08	
	A	0	8	13	.29	.17	1.1	.07	
	D	1	13	22	.55	.10	.8	.29	

The ability of each respondent to answer questions could be analyzed from the following dichotomous curve (see Figure 4). Based on this curve, it could be seen that respondents who had a low level of person item measure, then the probability of obtaining a score of 0 was higher. In contrast, respondents who had a higher level of person item measure tended to get a score of 1.

```
          DICHOTOMOUS CURVES
P      -+--------------+--------------+--------------+--------------+-
R  1.0 +                                                              +
O      |                                                              |
B      |0                                                            1|
A      | 000000                                                111111 |
B   .8 +       00000                                      11111       +
I      |            0000                              1111            |
L      |                0000                      1111                |
I      |                    000                111                    |
T   .6 +                       000          111                       +
Y      |                          000    111                          |
    .5 +                             ***                              +
O      |                          111    000                          |
F   .4 +                       111          000                       +
       |                    111                000                    |
R      |                1111                      0000                |
E      |            1111                              0000            |
S   .2 +       11111                                      00000       +
P      | 111111                                                000000 |
O      |1                                                            0|
N      |                                                              |
S   .0 +                                                              +
E      -+--------------+--------------+--------------+--------------+-
       -2             -1              0              1              2
         Person [MINUS] Item MEASURE
```

Figure 4. Dichotomous Curve

Pedagogy Competence

The results of the item analysis on the pedagogic competence showed that the reliability value of the items was .93, which was classified as very good (see Table 7). The high value of item reliability indicated that the questions developed were very consistent in measuring the ability of test participants.

Table 7. Summary of Measured Items

	Total Score	Count	Measure	Model Error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	27.1	60.0	.00	.34	1.00	.0	1.11	.0
SD	13.4	.0	1.41	.15	.13	1.1	.53	1.2
Max.	56.0	60.0	4.01	1.02	1.34	2.7	3.04	2.4
Min.	1.0	60.0	-3.24	.28	.28	-2.1	.60	-1.9

Real RMSE. .38 True SD. 1.36 Separation. 3.58 Reliability .93
Model RMSE .37 True SD. 1.36 Separation 3.66 Reliability .93
SE of Item Mean = .26

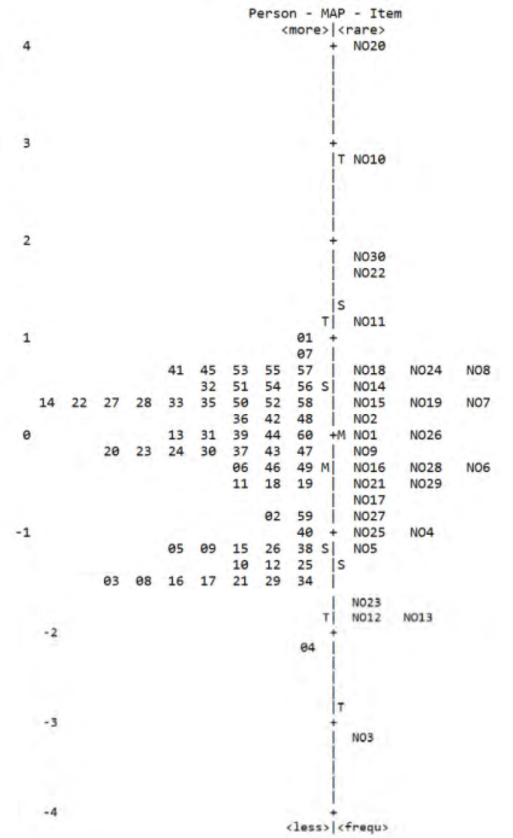


Figure 5. Wright's Map of Pedagogic Competence

Figure 5 refers to Wright's map that compared the respondents with the pedagogic competence of the questions. Based on the Wright map above, the most difficult question was question 20. Meanwhile, the lower the position of the item, the easier the question, such as question 3. Questions 20, 10, 30, and 22 could not be answered correctly by the respondent who had the highest score (i.e., respondent 01). The measured value of each item could be observed in the item measure table (see Figure 9).

Table 8. Measured Items of Pedagogic Competence

Entry Number	Total Score	Total Count	Measure	Model Se	Infit Mnsq	Infit Zstd	Outfit Mnsq	Outfit Zstd	Pt-Measure Corr.	Pt-Measure Exp	Exact Match Obs%	Exact Match Exp%	Item
20	1	60	4.01	1.02	1.04	.4	2.90	1.4	-.13	.08	98.3	98.3	20
10	3	60	2.86	.60	1.07	.3	3.04	2.1	-.12	.14	95.0	95.0	10
30	7	60	1.91	.41	1.00	.1	.87	-.2	.22	.21	88.3	88.3	30
22	9	60	1.60	.37	1.02	.2	.95	.0	.22	.23	85.0	85.0	22
11	12	60	1.23	.34	.86	-.7	.74	-.9	.45	.26	80.0	80.0	11
8	18	60	.64	.30	.87	-1.0	.76	-1.3	.48	.31	75.0	70.6	8
18	18	60	.64	.30	1.14	1.2	1.28	1.4	.10	.31	71.7	70.6	18
24	18	60	.64	.30	.95	-.4	.88	-.6	.38	.31	68.3	70.6	24
14	19	60	.55	.29	.90	-.9	.77	-1.3	.47	.31	66.7	69.5	14
15	21	60	.38	.29	.95	-.4	.91	-.5	.39	.32	66.7	67.4	15
19	21	60	.38	.29	.99	-.1	.98	-.1	.34	.33	70.0	67.4	19
7	22	60	.30	.28	1.06	.6	1.11	-.8	.25	.33	65.0	66.5	7
2	24	60	.14	.28	.95	-.5	.92	-.6	.41	.33	70.0	65.7	2
26	25	60	.07	.28	1.08	.8	1.05	.5	.26	.34	56.7	65.4	26
1	26	60	-.01	.28	.96	-.5	.91	-.7	.41	.34	65.0	65.1	1

Table 8. Continued

Entry Number	Total Score	Total Count	Measure	Model Se	Infit		Outfit		Pt-Measure		Exact Match		Item
					Mnsq	Zstd	Mnsq	Zstd	Corr.	Exp	Obs%	Exp%	
9	27	60	-.09	.28	.98	-.2	.96	-.3	.38	.35	61.7	65.0	9
28	30	60	-.32	.28	.84	-1.8	.81	-1.9	.55	.36	78.3	65.4	28
6	31	60	-.39	.28	1.00	.0	1.03	.3	.35	.36	66.7	65.7	6
16	31	60	-.39	.28	1.07	.8	1.08	.8	.27	.36	60.0	65.7	16
21	32	60	-.47	.28	.97	-.3	.96	-.4	.40	.36	70.0	66.2	21
29	32	60	-.47	.28	.88	-1.3	.85	-2.4	.51	.36	70.0	66,2	29
17	34	60	-.63	.29	.81	-2.0	.81	-1.8	.58	.36	78.3	67.2	17
27	36	60	-.78	.29	.78	-2.1	.78	-1.9	.61	.36	81.7	68.3	27
25	38	60	-.95	.29	1.34	2.7	1.36	2.4	-.04	.35	51.7	69.6	25
4	39	60	-1.03	.29	1.19	1.6	1.26	1.7	.11	.34	61.7	70.2	4
5	40	60	-1.11	.32	1.22	1.7	1.22	1.4	.10	.32	56.7	70.8	5
23	46	60	-1.90	.34	1.19	1.1	1.32	1.3	.07	.31	78.3	77.0	23
12	48	60	-1.90	.34	.95	-.2	.86	-.4	.38	.31	81.7	80.2	12
13	48	60	-1.90	.34	1.06	.4	1.20	.8	.21	.31	81.7	80.2	13
3	58	60	-3.24	.53	.93	.0	.60	-.6	.33	.20	93.3	93.3	3
MEAN	27.1	60.0	.00	.34	1.11	.0	1.11	.0			73.1	73.2	
SD	13.4	.0	1.41	.15	.53	1.1	.53	1.2			11.6	9.7	

Table 8 shows the measured values for each pedagogical competence of the question which was the level of difficulty of the questions. The question grouping category was based on the logit value of each item compared to the Standard Deviation (SD) value. The SD value in the calculation was 1.41. Question 20 was the most difficult question on this test with a logit value of 4.01, which was much greater than the SD score and was proven by the score obtained of 1, which meant that the question was only answered correctly by one respondent from the total number of respondents. Questions 10, 30, and 22 were difficult questions because the measured value was less than SD and had a positive value.

Table 9. Misfit Items of Pedagogic Questions

Entry Number	Total Score	Total Count	Measure	Model Se	Infit		Outfit		Pt-Measure		Exact Match		Item
					Mnsq	Zstd	Mnsq	Zstd	Corr.	Exp	Obs%	Exp%	
10	3	60	2.86	.60	1.07	.3	3.04	2.1	A-.12	.14	95.0	95.0	10
20	1	60	4.01	1.02	1.04	.4	2.90	1.4	B-.13	.08	98.3	98.3	20
25	38	60	-.95	.29	1.34	2.7	1.36	2.4	C-.04	.36	51.7	69.6	25
23	46	60	-1.68	.32	1.19	1.1	1.32	1.3	D .07	.32	78.3	77.0	23
18	18	60	.64	.30	1.14	1.2	1.28	1.4	E .10	.31	71.7	70.0	18
4	39	60	-1.03	.29	1.19	1.6	1.26	1.7	F .11	.35	61.7	70.2	4
5	40	60	-1.11	.29	1.22	1.7	1.22	1.4	G .10	.35	56.7	70.8	5
13	48	60	-1.90	.34	1.06	.4	1.20	.8	H .21	.31	81.7	80.2	13
7	22	60	.30	.28	1.06	.6	1.11	.8	I .25	.33	65.0	66.5	7
16	31	60	-.39	.28	1.07	.8	1.08	.8	J .27	.36	60.0	65.7	16
26	25	60	.07	.28	1.08	.8	1.05	.5	K .26	.34	56.7	65.4	26
6	31	60	-.39	.28	1.00	.0	1.03	.3	L .35	.36	66.7	65.7	6
22	9	60	1.60	.37	1.02	.2	.95	.0	M .22	.23	85.0	85.0	22
30	7	60	1.91	.41	1.00	.1	.87	-.2	N .22	.23	88.3	88.0	30
19	21	60	.38	.29	.99	-.1	.98	-.1	O .34	.33	70.0	67.4	19
9	27	60	.09	.28	.98	-.2	.96	-.3	O .38	.35	61.7	65.0	9
21	32	60	-.47	.28	.97	-.3	.96	-.5	N .40	.31	70.0	66.2	21
1	26	60	-.01	.28	.96	-.5	.96	-.2	M .41	.33	65.0	65.1	1
12	48	60	-1.90	.28	.95	-.2	.91	-.7	L .38	.31	81.7	80.2	12
2	24	60	.14	.28	.95	-.5	.86	-.4	K .41	.33	70.0	65.7	2
24	18	60	.64	.30	.95	-.4	.92	-.6	J .38	.20	68.0	70.6	24
15	21	60	.38	.53	.95	-.4	.88	-.6	I .39	.32	66.7	67.4	15
3	56	60	-3.24	.29	.93	.0	.91	-.5	H .33	.36	93.3	93.3	3
14	19	60	.55	.53	.90	-.9	.60	-.6	G .47	.31	66.7	69.5	14
29	32	60	-.47	.29	.88	-1.3	.77	-1.3	F .51	.26	70.0	66.2	29

Table 9. Continued

Entry Number	Total Score	Total Count	Measure	Model Se	Infit Mnsq	Infit Zstd	Outfit Mnsq	Outfit Zstd	Pt-Measure Corr.	Pt-Measure Exp	Exact Match Obs%	Exact Match Exp%	Item
8	18	60	.64	.28	.87	-1.0	.85	-1.4	E .48	.36	75.0	70.6	8
11	12	60	1.23	.30	.86	-.7	.76	-1.3	D .45	.36	80.0	80.0	11
28	30	60	-.32	.34	.84	-1.8	.74	-.9	C .55	.36	78.3	65.4	28
17	34	60	-.63	.28	.81	-2.0	.81	-1.9	B .58	.36	78.3	67.2	17
27	36	60	-.78	.28	.78	-2.1	.78	-1.9	A .61	.36	81.7	68.3	27
MEAN	27.1	60.0	.00	.34	1.00	.0	1.11	.0			73.1	73.2	
SD	13.4	.0	1.41	1.41	.13	1.1	.53	1.2			11.6	9.7	

Table 9 was a table of items fit order that showed the suitability of the items with the Rasch model. Table 8 could be used an indicator of whether an item needed to be re-evaluated or not. The item fit analysis was done by looking at the mean square (MNSQ) outfit value, the Z-standard outfit value, and the point measure correlation (PT-Measure Cor.) value. Based on the table above, question 10 got an MNSQ outfit value of more than 1.5 and a ZSTD value of more than 2 so the question could be said to be unfit with the Rasch model. Question 10 could be revised or not included in the next test. Question 20 got an MNSQ outfit score greater than 1.5 but the ZSTD score still met the standard, because the question could still be maintained with a few revisions. Question 25 got an MNSQ outfit score of less than 1.5 but the ZSTD outfit score did not meet the standard, therefore the question was said to be unfit.

The distractor analysis could be seen from the average ability value (see Table 10). If the trend increased in each answer option, the distractor worked well. Questions 10, 20, 25, and 23 showed an inconsistent order of average ability values, which meant the distractor option did not work well. In Figure 10, it could be seen that option A in question 10 was at the top with an average ability value of -.60, and option C which became the correct answer was at the bottom with an average ability value of -.73. This indicated a declining trend and the distractor on the matter was not good. The minus value on average ability indicated that there were respondents who had low abilities who could answer the correct answer so it could be said that the respondent only guessed correctly. Question 22 was the example of whose distractors worked well. In question 22, it was known that option A was in first place with an average ability value of -.95, and option C which became the correct answer was at the bottom with an average ability value of .08 so the trend of the average ability value increased, then the distractors were considered working well.

Table 10. The Distribution of Distractors for Each Item on the Pedagogic Competence

Entry Number	Data Code	Score Value	Data Count	Data %	Average Ability	S.E. Mean	Outfit MNSQ	PTMEA Corr.	Item
10 A	A	0	6	10	-.60	.31	.8	-.11	10
	D	0	50	83	-.29	.11	1.1	.11	
	E	0	1	2	.48		1.8	.13	
	C	1	3	5	-.73*	.68	3.1	-.12	
20 B	E	0	33	55	-.52	.14	.9	-.27	20
	B	0	6	10	-.23	.23	1.0	.04	
	D	0	18	30	-.06	.17	1.2	.22	
	A	0	2	3	.48	.18	1.8	.19	
	C	1	1	2	-1.14*		2.9	-.13	
25 C	D	0	1	2	-1.58		.3	-.21	25
	E	0	21	35	-.23	.13	1.4	.09	
	C	1	38	63	-.35*	.14	1.4	-.04	
23 D	C	0	1	2	-1.14		.5	-.13	23
	B	0	6	10	-.77	.41	1.1	-.19	
	E	0	7	12	-.05	.20	1.7	.13	
	A	1	46	77	-.30*	.12	1.1	.07	
22 M	A	0	2	3	-.95	.19	.4	-.15	22
	E	0	42	70	-.42	.13	1.0	-.16	
	B	0	7	12	-.19	.27	1.1	.07	
	C	1	9	15	.08	.18	.9	.22	

```
        DICHOTOMOUS CURVES
P      -+-------------+-------------+-------------+-------------+-
R  1.0 +                                                          +
O      |                                                          |
B      |0                                                        1|
A      | 000000                                           111111 |
B   .8 +       00000                                 11111        +
I      |            0000                         1111             |
L      |                0000                 1111                 |
I      |                    000           111                     |
T   .6 +                       000     111                        +
Y      |                          000 111                          |
    .5 +                             ***                           +
O      |                          111 000                          |
F   .4 +                       111     000                        +
       |                    111           000                     |
R      |                1111                 0000                 |
E      |            1111                         0000             |
S   .2 +       11111                                 00000        +
P      | 111111                                           000000 |
O      |1                                                        0|
N      |                                                          |
S   .0 +                                                          +
E      -+-------------+-------------+-------------+-------------+-
       -2            -1             0             1             2
        Person [MINUS] Item MEASURE
```

Figure 11. Dichotomous Curve

Based on Figure 11, the dichotomous curve explained the probability of the respondent's score revealed from the person item measure. Respondents who had a low person item measure were likely to get a higher score of 0. Meanwhile, respondents who had a high person item measure tended to get a score of 1.

Discussion

The development of teacher competency evaluation is one of the processes that need to be carried out to determine the readiness of teachers or pre-service teachers to face the learning process. Based on PP No. 74 of 2008, there are four teacher competencies, namely pedagogic, professional, social, and personality competencies. Professional competence is the ability of teachers to master learning materials in depth both theoretically and practically. Professional competence is the first aspect that a teacher must possess because it reflects how responsible the teacher deals with his works (Yurdakul, 2011). Pedagogic competence is the ability of teachers to convey their knowledge to students in accordance with learning theories and the development of science and technology. A teacher with good pedagogic competence will be able to teach the values of a life well. The progress of science and technology has now reached the era of the industrial revolution 4.0 so that every individual, including teachers, should be able to have competencies in accordance with the industrial revolution 4.0 in collaboration with the basic competencies of teachers. Pre-service teachers are also expected to have these competencies when developing the learning process in their near future careers. In this study, science teachers' competency evaluation has been developed in professional and pedagogic competencies according to important skills needed in the industrial revolution 4.0.

30 questions about evaluating the competence of pre-service science teachers in the professional and pedagogic competencies have been developed with a very good result of reliability. The reliability score is obtained through Rasch modeling using the Winsteps application. The high value of reliability means that the questions developed are very consistent to measure something assessed, in this study, namely the competence of pre-service science teachers. In addition, the question can be trusted to assess the criteria that have been set according to the question development grid (Amalia & Widayati, 2012). The higher the level of reliability implies that respondents who have high abilities will get a high score too, and vice versa. The reliability score is also one of the benchmarks for a test; whether someone was said to be good or not. Thus, this test is a good tool to measure the ability of pre-service teachers especially in professional and pedagogic competencies in accordance with the capabilities of the industrial revolution 4.0.

Based on the analysis of each item using the Rasch model, data were in the form of the respondent's logit value and professional and pedagogic test items, the level of difficulty of each item, the suitability of the item with the Rasch model, a description of the distractors for each test item, and a curve that describes the probability of the score obtained by the respondents based on their abilities. The results of the item analysis show that there are some difficult questions because only a few respondents can answer correctly and meet the standard criteria for difficult and very difficult questions (measure value > SD value). The questions developed, both professional and pedagogical, show that most of the questions are moderate, meaning that they are not too difficult and not too easy. More than half of the total respondents can answer these questions. The most difficult questions in professional competence are questions that measure complex problem solving and critical thinking skills on the material of the influence of the moon's revolution. In this question, it is possible that respondents have difficulty choosing answers to questions that look homogeneous. The homogeneity of the answer options is one component that can be used to increase the level of difficulty of the questions (Applegate et al., 2019). The more similar the answer options, the more misleading the respondent will be, on the contrary, if the answer options are very different from the correct answer (heterogeneous), then the question is easy to guess. Therefore, it is also very important to consider distractors in each item (Gierl et al., 2017). Based on the results of the study, questions that are classified as difficult have answer options that do not deceive or the distractor does not work on the question.

A reliable test can be analyzed for its level of validity (Fraenkel et al., 2012). The validity of the test instrument is the suitability, correctness, and usefulness of each conclusion that is interpreted based on the results of data analysis from the instrument used. The test questions that have been developed are theoretically validated and declared valid. In addition, empirical validation was also carried out based on the analysis of respondents' answers using the Rasch model. The validity analysis using the Rasch model can be observed from the data item misfit which shows how fitted the questions developed with the Rasch model are. There are two statistical calculations used to determine the suitability of the test with the Rasch model.

Conclusion

This study has produced a competency evaluation instrument for science teacher candidates in the industrial revolution 4.0 era as many as 30 multiple-choice questions in the professional and pedagogic competencies and has been tested on 60 pre-service teachers majoring in science education at two State universities in Indonesia. The item analysis was carried out using the Rasch model and resulted in the reliability value of the professional competence test items of .95 which was classified as special and the pedagogic competence test items were stated to be very good with a reliability value of .93. There was one question, both in the professional and pedagogic competencies, which did not fit the Rasch model, while the other items were acceptable. Therefore, the test instrument is said to be valid so that it can measure the competence of prospective science teachers in the 4.0 industrial revolution era.

Recommendations

The present study suggests that future research can develop competency evaluation for different fields, including language education. It is really important to note that those fields require specific evaluation of the pedagogical and professional aspects. For instance, the competence of teaching is indeed different from the competence of having good skills in languages, which become parts of professionalism. This evaluation is necessary to maintain the quality of pre-service teachers, not only in the science field but also in others such as the language fields. In coping with the use of the competency evaluation, the practitioners can refer to this competency evaluation for the science field. If they come from different branches, they can use the evaluation sheet as a reference to develop another. By having competency evaluation for various fields, the pre-service teachers' professional and pedagogic competencies can be guaranteed before they are graduated from the teacher training institution. Future research might do the same evaluation sheet development to evaluate pre-service language teachers (i.e., English and Chinese) since there will be so more complex professional aspects, including but not limited to conceiving good language skills.

Limitations

The present study is only limited to the science pre-service teachers in Indonesia. The developed competency evaluation might not be applicable for other countries since the curriculum must be dissimilar. Moreover, the present study's participants are only limited to two State universities in Java Island, thus, the final judgment can be generalized as whole Indonesian pre-service teachers' cases. Therefore, there are still many things to be developed and further undertaken due to this limitation.

Acknowledgments

The authors thank Universitas Negeri Surabaya, Universitas Negeri Yogyakarta, Universitas Pendidikan Indonesia, Universitas Pendidikan Ganesha, dan Universitas Negeri Gorontalo for the endless support to accomplish this collaborative works. They also thank science pre-service teachers who participated in the present study.

Funding

DRPM (Directorate of Research and Community Service) through the Leading Research Consortium for Higher Education in the field of Education

Authorship Contribution Statement

Susantini: Initiating ideas, designing conceptual works, and planning and leading the research. Kartowagiran, Hamdi, and Jaedun: Drawing the research backgrounds and methods and conducting the data collection and data analysis. Wesnawa: Collecting reviews of literatures, collecting data, and disseminating data. Sunendar: Collecting reviews of litearatures, collecting data, and writing results. Laliyo: Writing discussion, conclusion, recommendations, limitations, acknowledgments, funding, and references, and assisting other authors.

References

Abdullah, Q. A., Humaidi, N., & Shahrom, M. (2020). Industry revolution 4.0: The readiness of graduates of higher education institutions for fulfilling job demands. *Romanian Journal of Information Technology and Automatic Control*, *30*(2), 15-26. https://doi.org/10.31004/edukatif.v3i5.1055

Akimova, L., Akimov, O., Mihus, I., Koval, Y., & Dmitrenko, V. (2020). Improvement of the methodological approach to assessing the impact of public governance on ensuring the economic security of the state. *Financial and Credit Activity: Problems of Theory and Practice*, *4*(35), 180-190. https://doi.org/10.18371/fcaptp.v4i35.221969

Alda, R., Boholano, H., & Dayagbil, F. (2020). Teacher education institutions in the Philippines towards education 4.0. *International Journal of Learning, Teaching and Educational Research*, *19*(8), 137-154. https://doi.org/10.26803/ijlter.19.8.8

Amalia, A. N., & Widayati, A. (2012). Analisis butir soal tes kendali mutu kelas XII SMA mata pelajaran ekonomi akuntansi di Kota Yogyakarta tahun 2012 [Analysis of quality control test items for class XII SMA economics accounting subject in Yogyakarta City in 2012]. *Jurnal Pendidikan Akuntansi Indonesia*, *10*(1), 1-26. https://doi.org/10.21831/jpai.v10i1.919

Applegate, G. M., Sutherland, K. A., Becker, K. A., & Luo, X. (2019). The effect of option homogeneity in multiple-choice items. *Applied Psychological Measurement*, *43*(2), 113-124. https://doi.org/10.1177/0146621618770803

Arifin, Z. (2009). *Evaluasi pembelajaran* [Learning evaluatio]. Remaja Rosdakarya.

Bakar, N. S. A., Maat, S. M., & Rosli, R. (2020). Mathematics teacher's self-efficacy of technology integration and technological pedagogical content knowledge. *Journal on Mathematics Education*, *11*(2), 259-276. http://doi.org/10.22342/jme.11.2.10818.259-276.

Ellitan, L. (2020). Competing in the era of industrial revolution 4.0 and society 5.0. *Jurnal Maksipreneur: Manajemen, Koperasi, dan Entrepreneurship*, *10*(1), 1-12. https://doi.org/10.30588/jmp.v10i1.657

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education. McGraw*-Hill.

Gierl, M. J., Bulut, O., Guo, Q., & Zhang, X. (2017). Developing, analyzing, and using distractors for multiple-choice tests in education: A comprehensive review. *Review of Educational Research*, *87*(6), 1082–1116. https://doi.org/10.3102/0034654317726529

Gleason, N. W. (2018). *Higher education in the era of the fourth industrial revolution*. Springer Nature. https://doi.org/10.1007/978-981-13-0194-0

Haladyna, T. M., Rodriguez, M. C., & Stevens, C. (2019). Are multiple-choice items too fat? *Applied Measurement in Education*, *32*(4), 350-364. https://doi.org/10.1080/08957347.2019.1660348

Indira, E. W. M., Hermanto, A., & Pramono, S. E. (2020, June). Improvement of teacher competence in the industrial revolution era 4.0. In S. Xiao (Eds.), *Proceedings of the International Conference on Science and Education and Technology (ISET 2019)* (pp. 350-352). Atlantis Press. https://doi.org/10.2991/assehr.k.200620.068

Juhji, J., & Nuangchalerm, P. (2020). Interaction between scientific attitudes and science process skills toward technological pedagogical content knowledge. *Journal for the Education of Gifted Young Scientists*, *8*(1), 1-16. https://doi.org/10.17478/jegys.2020.XX

Kopotun, I. M., Durdynets, M. Y., Teremtsova, N. V., Markina, L. L., & Prisnyakova, L. M. (2020). The use of smart technologies in the professional training of students of the Law Departments for the development of their critical thinking. *International Journal of Learning, Teaching and Educational Research*, *19*(3), 174-187. https://doi.org/10.26803/ijlter.19.3.10

Macaro, E., Akincioglu, M., & Han, S. (2020). English medium instruction in higher education: Teacher perspectives on professional development and certification. *International Journal of Applied Linguistics*, *30*(1), 144-157. https://doi.org/10.1111/ijal.12272

Maryanti, N., Rohana, R., & Kristiawan, M. (2020). The principal's strategy in preparing students ready to face the industrial revolution 4.0. *International Journal of Educational Review*, *2*(1), 54-69. https://bit.ly/3L0rnV7

Muryadi, A. D. (2017). Model evaluasi program dalam penelitian evaluasi [Program evaluation model in evaluation research]. *Jurnal Ilmiah Penjas*, *3*(1). https://bit.ly/39HG02E

Nuryanti, S., Masykuri, M., & Susilowati, E. (2018). Analisis Iteman dan model Rasch pada pengembangan instrumen kemampuan berpikir kritis peserta didik sekolah menengah kejuruan [Iteman's analysis and Rasch's model on the development of critical thinking skills of vocational high school students]. *Jurnal Inovasi Pendidikan IPA*, *4*(2), 224-233. https://doi.org/10.21831/jipi.v4i2.21442

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. Routledge. https://doi.org/10.4324/9781003117452

Philbeck, T. D. (2017). Can we build a resilient employment market for an uncertain future? In LaGrandeur, K., Hughes, J. *Surviving the Machine Age* (pp. 83-106). Palgrave Macmillan. https://doi.org/10.1007/978-3-319-51165-8_6

Rahiem, M. D. (2020). The emergency remote learning experience of university students in Indonesia amidst the COVID-19 crisis. *International Journal of Learning, Teaching and Educational Research*, *19*(6), 1-26. https://doi.org/10.26803/ijlter.19.6.1

Türkoguz, S. (2020). Investigation of three-tier diagnostic and multiple choice tests on chemistry concepts with response change behaviour. *International Education Studies*, *13*(9), 10-22. https://doi.org/10.5539/ies.v13n9p10

Wahyuni, D. (2018). Peningkatan kompetensi guru menuju era revolusi industri 4.0 [Improving teacher competence towards the industrial revolution era 4.0]. *Kajian Singkat Terhadap Isu Aktual dan Strategis Puslit Badan Keahlian DPR RI*, *X* (24), 13-18. https://bit.ly/3FyrZAk

Wartoni, W., & Benyamin, P. I. (2020). Strategi pengembangan tes objektif (pilihan ganda) [Objective test development strategy (multiple choice)]. *Diegesis: Jurnal Teologi*, *5*(1), 33-53. https://doi.org/10.46933/DGS.vol5i1%25p

World Economic Forum. (2016, January 1). *The future of jobs*. https://bit.ly/3FBIyvl

Xu, M., David, J. M., & Kim, S. H. (2018). The fourth industrial revolution: Opportunities and challenges. *International Journal of Financial Research*, *9*(2), 90-95. https://doi.org/10.5430/ijfr.v9n2p90

Yurdakul, I. K. (2011). An evaluative case study on professional competency of preservice information technology teachers. *Turkish Online Journal of Educational Technology, 10*(3), 33-53. https://bit.ly/39TIYkS

Developing Competency Evaluation of Pre-service Science Teachers in Industrial Revolution 4.0: Revealing Pedagogic and Professional Competencies

ORIGINALITY REPORT

13%	10%	10%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	jurnal.univpgri-palembang.ac.id Internet Source	1%
2	E Susilaningsih, M Nuswowati, N Aprilia, A Luthfiyah. "Dissemination of test instruments as product of the development research to measure the problem-solving ability of class X students by online in the pandemic period", Journal of Physics: Conference Series, 2021 Publication	1%
3	digitalcommons.unl.edu Internet Source	1%
4	Naicker, Aloshna. "An Investigation of the Quality of the Test Items on the Numerical Scales of the isiZulu Translated Version of the Junior South African Intelligence Scales (Jsais) (G-IQ8)", University of Johannesburg (South Africa), 2021 Publication	<1%

www.semanticscholar.org

5	Internet Source	<1%
6	Chatchawal Sarapoom. "Cooperative Practices to Enhance the Quality of Work-Integrated Learning at Nong Khai Technical College", International Journal of Educational Methodology, 2021 Publication	<1%
7	Submitted to Universitas Pendidikan Ganesha Student Paper	<1%
8	Edi Subarkah, Badrun Kartowagiran, Sumarno Sumarno, Syukrul Hamdi*, Abdul Rahim. "The Development of Life Skill Education Evaluation Model at Life Skill Training Centre", International Journal of Educational Methodology, 2022 Publication	<1%
9	Connie Connie, Eko Risdianto, Meizul Zuki, Adi Asmara. "Implementation of E-Learning in Learning Management Education Innovation During The Covid-19 Pandemic", Tadbir : Jurnal Studi Manajemen Pendidikan, 2021 Publication	<1%
10	Submitted to The Hong Kong Institute of Education Student Paper	<1%
11	www.ijiet.org	

Internet Source

<1 %

12 uia.e-journal.id
Internet Source

<1 %

13 www.nspe.org
Internet Source

<1 %

14 Samsul Huda, M Syariffudien Zuhrie, I Gusti Putu Asto Buditjahjanto, Luthfiyah Nurlaela. "Implementation of Educational Robotic into Teaching-Learning Process to Enhance Students Skills in the Science and Technology", Journal of Physics: Conference Series, 2021
Publication

<1 %

15 eprints.uny.ac.id
Internet Source

<1 %

16 ijmmu.com
Internet Source

<1 %

17 journal.um-surabaya.ac.id
Internet Source

<1 %

18 Zulkefli Mansor, Saadiah Yahya, Noor Habibah Arshad. "Empirical Study of Cost Management Success Determinants in Agile based Software Development Project: A Rasch Measurement Model Analysis", Procedia - Social and Behavioral Sciences, 2013

<1 %

Publication

19	pt.scribd.com Internet Source	<1 %
20	Dewi Ayu, Murni Ramli, Joko Ariyanto. "Item Analysis of Learning Progression Test of Carbon Transformation Process", Proceedings of the 4th International Conference on Learning Innovation and Quality Education, 2020 Publication	<1 %
21	Rosane Medriati, Eko Risdianto, Andik Purwanto, Kusen Kusen. "Rasch Model Analysis on the Development of Digital Learning Model Using MOOCs in Practical Courses at University", AL-ISHLAH: Jurnal Pendidikan, 2022 Publication	<1 %
22	files.eric.ed.gov Internet Source	<1 %
23	Submitted to Universitas Islam Bandung Student Paper	<1 %
24	eprints.utm.my Internet Source	<1 %
25	Submitted to Barnet and Southgate College Student Paper	<1 %
26	medium.com Internet Source	

<1 %

27	theease.org Internet Source	<1 %
28	www.ukm.my Internet Source	<1 %
29	Rupal K Shah, Meenu Saraf. "Exploring Inorganic Phosphate Solubilizing Trait of Halotolerant Rhizobacteria Isolated From Cuminum cyminum", International Journal of Research in Advent Technology, 2019 Publication	<1 %
30	consortiacademia.org Internet Source	<1 %
31	hdl.handle.net Internet Source	<1 %
32	ojs.fkip.ummetro.ac.id Internet Source	<1 %
33	Siti Aminah Osman, Syahdatul Isnain Naam, Othman Jaafar, Wan Hamidon Wan Badaruzzaman, Riza Atiq Abdullah O.K. Rahmat. "Application of Rasch Model in Measuring Students' Performance in Civil Engineering Design II Course", Procedia - Social and Behavioral Sciences, 2012 Publication	<1 %

34	tribune.net.ph Internet Source	<1 %
35	ejournal.undiksha.ac.id Internet Source	<1 %
36	staffnew.uny.ac.id Internet Source	<1 %
37	Aziz, Azrilah Abd., Nuraini Khatimin, Azami Zaharim, and Tuan Salwani Awang Salleh. "Evaluating multiple choice items in determining quality of test", Proceedings of 2013 IEEE International Conference on Teaching Assessment and Learning for Engineering (TALE), 2013. Publication	<1 %
38	Kusen Kusen. "Rasch Test for Digital Learning Material Development Based on MOOCS", Tadbir : Jurnal Studi Manajemen Pendidikan, 2022 Publication	<1 %
39	Mona Isa, Mazlan Abu Bakar, Mohamad Sufian Hasim, Mohd Khairul Anuar, Ibrahim Sipan, Mohd Zali Mohd Nor. "Data quality control for survey instrument of office investors in rationalising green office building investment in Kuala Lumpur by the application of Rasch analysis", Facilities, 2017 Publication	<1 %

40	cwww.intechopen.com Internet Source	<1%
41	ejournal.unib.ac.id Internet Source	<1%
42	www.hawaii.edu Internet Source	<1%
43	www.scirp.org Internet Source	<1%
44	Agustin Kumala Dewi, Sentot Budi Rahardjo, Suryadi Budi Utomo. "Item Analysis of an Instrument to Measure Chemistry Students' Critical Thinking Skills Based on Science, Technology, Engineering and Mathematics (STEM) on Solubility and Solubility Product (KSP) material using the Rasch model", Proceedings of the 4th International Conference on Learning Innovation and Quality Education, 2020 Publication	<1%
45	Dedy Triono, Riyanarto Sarno, Kelly R. Sungkono. "Item Analysis for Examination Test in the Postgraduate Student's Selection with Classical Test Theory and Rasch Measurement Model", 2020 International Seminar on Application for Technology of Information and Communication (iSemantic), 2020	<1%

Publication

46	Eko Risdianto, Wachidi Wachidi, Riyanto Riyanto, Alexon Alexon, Irwan Fathurrochman, Kusen Kusen. "Blended Learning Model Based on Massive Open Online Courses (MOOCs) Assisted by Augmented Reality (BMA) Model as the Electronic Learning Media in the Pandemic Covid-19", AL-ISHLAH: Jurnal Pendidikan, 2021 Publication	<1 %
47	Shalahudin Ismail, Ma'mun Zahrudin, Uus Ruswandi, Erihadiana Erihadiana. "The Competence of Millennial Islamic Education Teachers in Facing The Challenges of Industrial Revolution", Nazhruna: Jurnal Pendidikan Islam, 2020 Publication	<1 %
48	doaj.org Internet Source	<1 %
49	gtg.webhost.uoradea.ro Internet Source	<1 %
50	www.abacademies.org Internet Source	<1 %
51	www.e-iji.net Internet Source	<1 %
52	www.jurnal.unsyiah.ac.id Internet Source	<1 %

53	"English as the Medium of Instruction in Turkish Higher Education", Springer Science and Business Media LLC, 2022 Publication	<1 %
54	H Z Puspitaningrum, Wasis, T Prastowo. "High Order Thinking Skills Students Through Multi-Representation Test on Newtons Law Study", Journal of Physics: Conference Series, 2021 Publication	<1 %
55	Malek, Heather Leigh. "The Acquisition of English Prepositions in First Language Speakers of Northern Sotho and Afrikaans: A Cognitive Linguistic Study", University of Johannesburg (South Africa), 2021 Publication	<1 %
56	Melor Masdoki, Rosseni Din, Mohd Effendi @ Ewan Mohd. Matore. "Teaching 4.0 Competency in Higher Learning Institutions: A Systematic Mapping Review", International Journal of Learning, Teaching and Educational Research, 2021 Publication	<1 %
57	Rita Green. "Statistical Analyses for Language Testers", Springer Science and Business Media LLC, 2013 Publication	<1 %

58	Rui-Ni Wang, Yuan-Cheng Chang*. "Effect of Intrinsic Motivation on Junior High School Students' Creativity: Mediating Role of Cognitive Flexibility", International Journal of Educational Methodology, 2022 Publication	<1 %
59	Sharon Gondodiputro, Guswan Wiwaha, Melly Lionthina, Deni Kurniadi Sunjaya. "Reliability and validity of the Indonesian version of the World Health Organization quality of life-old (WHOQOL-OLD): a Rasch modeling", Medical Journal of Indonesia, 2021 Publication	<1 %
60	Toshikatsu Kaneda, Shinichi Takabatake, Yasuhiro Higashi, Yuka Horishima et al. "Evaluation of psychometric properties of the activities of daily living scale of motor function used by caregivers using Rasch analysis", Journal of Physical Therapy Science, 2020 Publication	<1 %
61	myscholar.umk.edu.my Internet Source	<1 %
62	repository.up.ac.za Internet Source	<1 %
63	www.scribd.com Internet Source	<1 %

64	Cao, Hui. "Development and Evaluation of the Authentic-Material-Based English Reading Test (AERT) as a Measure of 12th-Grade Students in the Chinese EFL Educational Context: A Rasch Analysis.", Niagara University, 2020 Publication	<1%
65	Di Terlizzi, Marisa Ann. "A Reading Intervention Programme for Adolescents", University of Johannesburg (South Africa), 2021 Publication	<1%
66	Mareana, Tsakani Ruth. "An Investigation of the Quality of Items on the Translated Sesotho Junior South African Individual Scale-GIQ8", University of Johannesburg (South Africa), 2021 Publication	<1%
67	Subash Balan, Zaida Nor Zainudin, Habibah Ab Jalil. "Understanding and Readiness in Facing IR 4.0 Future Skills Transformation among UPM Trainee Counsellors", Asian Social Science, 2021 Publication	<1%
68	Budi Utami, Sulistyo Saputro, Ashadi, Mohammad Masykuri. "Performance asessment to measure students' explanations	<1%

in chemistry learning", Journal of Physics: Conference Series, 2021
Publication

---

69 Dedi Kuswandi, Punaji Setyosari, Hutkemri Hutkemri, Nunung Suryati et al. "Development Of Life-Based Curriculum Model Designs In The Global Era", Proceeding on International Conference of Science Management Art Research Technology, 2020 <1%
Publication

---

70 Nurul Huda. "Experiences of Participants in Teacher Professional Education on Obtaining Soft Skills: A Case Study in Indonesia", European Journal of Educational Research, 2021 <1%
Publication

---

71 D. Karakhanyan, R. Kirschner. "Spinorial R operator and Algebraic Bethe Ansatz", Nuclear Physics B, 2020 <1%
Publication

---

72 Setiyani S., Waluya Y., Sukestiyarno Sukestiyarno, Adi Nur. "Mathematical Reflective Thinking Process of Prospective Elementary Teachers Review from the Disposition in Numerical Literacy Problems", International Journal of Educational Methodology, 2022 <1%
Publication

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		