How to develop test for measure critical and creative thinking skills of the 21st century skills in POPBL?

S N Muhajir*, S Utari and I R Suwarma
Program Studi Magister Pendidikan Fisika, Sekolah Pascasarjana Universitas Pendidikan Indonesia, Jl. Dr. Setiabudi No. 229, Bandung 40154, Indonesia

* sitinurdiantimuhajir@student.upi.edu

Abstract. 21st century skills are a set of capabilities students must possess to succeed in the information age. Critical and creative thinking skills are part of the basic skills in the 21st century addition of collaborating and communication skills. Therefore, we need a measuring instrument that can measure these skills. In this study, presented a way of developing tests to measure critical and creative thinking for the matter of direct current. The tests were developed, adapted to the Buck Institute’s critical and creative thinking abilities. Based on expert judgment and analysis using anates application, the test instruments were well developed and could be used to measure student’s critical and creative thinking. The findings show that the elusive test constructs greatly influence the validity of the test.

1. Introduction
Demands of work in this era poses new challenges for education systems around the world. Schools and teachers should prepare students to have skills to work on teams and use a variety of sciences as well as using technology for solving problems. Students must learn how to find and analyze information to create new products, new knowledge, and better services [1,2]. These skill that must be possessed by students to face this century was 21st century skills, which includes critical thinking skill, creative thinking skill, collaborative and communicating [3,4].

Based on the results of literature studies, 21st century skills are a very important skill possessed by humans. Much of the research done in efforts to develop 21st century skills include Fong which states that the skills of the century 21 is a matter of urgency for postgraduate students in Malaysia [5], besides according to (skill), 21st century skills need to be trained, as well as Turiman states that the skills of the century It is important to examine the need for human survival in the 21st century [6]. Husin states that the skills of 21st century students in Malaysia on the aspect of critical thinking and creative thinking are in the less category. But it is good in the aspects of collaboration and communicating [7]. Furthermore, suggested lesson to trained 21st century skills according Husin was POPBL or project oriented problem based learning.

Measuring 21st century skills of learners requires a written test assessment, in addition to being used to determine and exploring profile of learners, but can also be used as a means to train students ability to think at a higher level (critical and creative). In order for learners to answer the question, it requires high-level reasoning logical way of thinking high. High logical thinking is needed by learners in the face of problems in this era, especially problems that can be answered by using knowledge,
understanding, and skills they possess and connect them in new situations [8]. Thus, to measure 21st century skills we need a good instrument.

A good instrument is capable of measuring what it wants to measure, so that easily to interpret the true state of the student. We need to pay attention to know the quality of the test instrument. When test measures what to measure so that test have a good validity. Beside that, the test scores should be consistence after repeatedly tested [9-11]. Based on what has been described, important to develop good test instruments to provide an overview of students 21st century skills in POPBL.

2. Method
This research tries to develop a test instrument to measure the extent to which learners master the critical and creative thinking skills pre service physics teacher on Ohm's law, Kirchhoff's law and RC material. The item of the test instrument was an essay that contains current issues or problems. The method was used in this research was research and development method (R & D) Borg and Gall [12].

The 21st century skills skill development model (critical and creative thinking skills) is shown in Figure 1, in this study developed a student's critical and creative thinking skills test adapted to the 21st century skills framework for project based learning from the Buck Institute. Validation of tests was conducted by seeking the consideration of five physicists. Furthermore the test device that has fulfilled the content validity is tested in the field involving 24 students of mechatronics engineering study program at Polytechnic Manufacturing Bandung. The trial was held on March 15, 2018.

The test devices that have been tested are then revised and re-tested to a larger sample size, involving 72 pre service physics students. The second phase trial was held on 26 March 2018 at UIN Sunan Gunung Djati Bandung. The results of this test is done with the analysis of item, validity test and reliability test with the help of Anates application. The test tested amounted to eighteen grains covering three materials. The indicators used consisted of six indicators, consisting of three indicators of critical thinking skills and three indicators of creative thinking skills.

Figure 1. Scheme or the R&D Borg and Gall method.

Figure 1 shows development research methods that start from planning, product manufacturing, testing, and product results. This research adapts the development methods of Borg and Gall that are tailored to the needs of the study. Many development studies using this method include Fathimah and Redhana [13,14].
3. Result and discussion

3.1. Expert validation test result
Validity refers to the feasibility of a test used as a measuring tool. The developed instrument should be able to measure the aspect to be measured [13,11]. Validation involves five physicists who are lecturers at Universitas Pendidikan Indonesia. Experts provide the assessments, criticisms, and provide suggestions that can be used to improve the instruments developed. Experts provide an assessment of three aspects, namely aspects of materials, construction, and language selection. Validation results from five experts were: 1) the problem given is too long, the resistor image in problem 2 was not related to the problem and should be replaced with the LED lamp image, 2) the sentence for question number 1e, 2e, and 3e are replaced, 3) change the answer key of the question with indicators of critical and creative thinking. From the overall comments given by the experts, the developed test is considered to have met the standard of material and language aspects. Each expert is given a judgment sheet containing a checklist of eligibility questions on a scale A through D with eligibility criteria: (A) very feasible, (B) feasible, (C) is reasonable, and (D) is not feasible. Based on the assessment and input provided from the expert, all items fall into the category used with some revisions to the language aspect. The percentage of expert validation results is presented in Figure 2.

Figure 2. Percentage of expert validation results.

Figure 2 shows that from the 18 test items tested by the expert, 83% of the questions are categorized as feasible, while the remaining 17% are considered very feasible. This indicates that the test can be tested with little revision.

3.2. Legibility test result
The test of legibility is carried out by discussion activities between researchers and four students. This activity was conducted at Universitas Islam Negeri Sunan Gunung Djati. This method is chosen to make it easier to find the problem on each item. In this legibility test, the four students were asked to read, ask questions, criticize, and provide advice on the readability of the whole item. Then the matter discussed one by one to find students understanding of the item. The result of this legibility test is the replacement of some words to make the matter easy to understand. The word "tersusun" in problem 1 is replaced with the word "disusun". Then the word "kemukakan" on the questions no 1a, 2a and 3a replaced with the word "buatlah". In addition, students are also asked to provide criticism and suggestions. Student criticize that the problem given is too long, the student requests that the problems 1, 2, and 3 be shorter and clearer. This legibility test is important to do, because students understanding of the problem influence on student outcomes [13].

3.3. Trial result
Test validity is obtained from the data processing by using anates application. Table 1 shows the results of trials. The test reliability tested is 0.85 which is in the high category both limited and extensive trial result.
Table 1. Limited trial result.

| Question Number | Distinct power (%) | Difficulty level | Correlation | Validity       | Result |
|-----------------|--------------------|------------------|-------------|----------------|--------|
| 1a              | 31.58              | Medium           | 0.625       | Very Significant | Used   |
| 1b              | 22.81              | Medium           | 0.466       | Significant    | Used   |
| 1c              | 29.82              | Easy             | 0.650       | Very Significant | Used   |
| 1d              | 36.84              | Medium           | 0.622       | Very Significant | Used   |
| 1e              | 26.32              | Medium           | 0.554       | Very Significant | Used   |
| 1f              | 36.84              | Easy             | 0.791       | Very Significant | Used   |
| 2a              | 31.58              | Medium           | 0.373       | -              | Revised|
| 2b              | 26.32              | Medium           | 0.678       | Very Significant | Used   |
| 2c              | 7.02               | Medium           | 0.445       | Significant    | Used   |
| 2d              | 16.07              | Medium           | 0.689       | Very Significant | Used   |
| 2e              | 12.28              | Medium           | 0.459       | Significant    | Used   |
| 2f              | 16.67              | Medium           | 0.514       | Significant    | Used   |
| 3a              | 44.44              | Medium           | 0.689       | Very Significant | Used   |
| 3b              | 38.89              | Medium           | 0.750       | Very Significant | Used   |
| 3c              | 22.81              | Medium           | 0.678       | Very Significant | Used   |
| 3d              | 16.37              | Medium           | 0.540       | Very Significant | Used   |
| 3e              | 26.32              | Medium           | 0.678       | Very Significant | Used   |
| 3f              | 36.84              | Easy             | 0.633       | Very Significant | Used   |

Based on the results of limited trial data showed in table 1, there are two questions with low validity values from 18 question, therefore the question need to be revised. Those questions was number 2d and 3d. 2d question contains the command to give response to the data. The revision is to change the data writing ways in the table so that more readable. Similar to the 2d problem, the 3d question also contains commands for providing feedback based on data. The main problem about this 3d number caused by there are two tables presented, this makes the respondent confused to answer question no 3d. In addition to revising the 2d and 3d questions, researchers also conducted an analysis of the results of the answers to improve the rubric assessment. The entirety of the items that are prepared, also requires the students ability to apply their knowledge and reasoning [14,15]. For extensive trial result presented in table 2.

Table 2. Extensive trial result.

| Question Number | Distinct power (%) | Difficulty level | Correlation | Validity       | Result |
|-----------------|--------------------|------------------|-------------|----------------|--------|
| 1a              | 33.33              | Medium           | 0.689       | Very Significant | Used   |
| 1b              | 27.78              | Medium           | 0.539       | Significant    | Used   |
| 1c              | 38.89              | Easy             | 0.721       | Very Significant | Used   |
| 1d              | 33.33              | Medium           | 0.851       | Very Significant | Used   |
| 1e              | 44.44              | Medium           | 0.593       | Very Significant | Used   |
| 1f              | 44.44              | Easy             | 0.770       | Very Significant | Used   |
| 2a              | 33.33              | Medium           | 0.689       | Very Significant | Used   |
| 2b              | 27.78              | Medium           | 0.698       | Very Significant | Used   |
| 2c              | 16.67              | Medium           | 0.577       | Very Significant | Used   |
| 2d              | 16.67              | Medium           | 0.730       | Very Significant | Used   |
| 2e              | 33.33              | Medium           | 0.678       | Very Significant | Used   |
| 2f              | 44.44              | Medium           | 0.686       | Very Significant | Used   |
| 3a              | 44.44              | Medium           | 0.851       | Very Significant | Used   |
| 3b              | 38.89              | Medium           | 0.530       | Significant    | Used   |
| 3c              | 38.89              | Medium           | 0.690       | Very Significant | Used   |
| 3d              | 16.67              | Medium           | 0.570       | Very Significant | Used   |
| 3e              | 22.22              | Difficult        | 0.463       | Significant    | Used   |
| 3f              | 44.44              | Easy             | 0.638       | Very Significant | Used   |

Table 2 shows that the test developed is good. 14 of the 18 questions are in a very significant category, while the rest belong to a significant category. One of the 21st century learning that can develop critical thinking, creative, collaborative and communicative skills is problem based learning, project oriented
learning, and project oriented problem based learning [16,17]. However, even though the learning that has been done is trained in 21st century skills, it would be meaningless if the measuring instruments used to measure 21st century skills are invalid. Based on the results of the study, all the items developed can be used to measure the skills of 21st century students. The content of the developed test material is the result of experimental data in the laboratory.

Questions in tests of critical and creative thinking skills developed, adapted from the driving question developed by Buck Institute [18]. The developed test should include several categories that include good content, understood by students, in the form of open-ended questions, curiosity, interest, and real world implication [19]. Open questions allow students to come up with some direct questions. Thus, student thinking becomes directed [20]. The combination of direct questions and open questions will make it easier for students to understand key facts that encourage the use of various disciplines in problem solving.

4. Conclusion
21st century skills was an important skills for students to be trained with problem based learning or project oriented learning. The results of the analysis show that the test had been developed can be used to measure pre service physics teacher’s critical and creative thinking skills. In the development of the tests, open-ended questioning was the most important factor for generating student’s critical and creative thinking skills, especially in POPBL learning. In addition to open questions, language selection in real world application issues also influence the test quality. Instrument critical and creative thinking skills should be tested after students entered the 21st century learning course, for example the application of project oriented problem based learning.

Acknowledgments
We thank our colleagues from Universitas Islam Sunan Gunung Djati Bandung who provided insight and expertise that greatly assisted the research. Thanks to experts who have provided feedback, and the laboratory assistant in Basic Physics Laboratory of the Universitas Pendidikan Indonesia, who has assisted the experimental process for research purposes. We would also like to show our gratitude to the librarian of the physics engineering faculty of Institut Teknologi Bandung who has provided an opportunity and assist in finding sources of information for research purposes.

References
[1] Silva E Y, Zulkardi Z and Darmawijoyo D 2011 J. Pend. Matematika 5 1
[2] Griffin P, Care E and McGaw B 2012 J. Springer 1 15
[3] Voogt J and Roblin N P 2012 J. Curr. Stydys 44 3
[4] Voogt J, Erstad O, Dede C and Mishra P 2013 J. of computer assisted learning 29 5
[5] Fong L, Sidhu G and Fook C Y 2014 Procedia-Social and Behavioral Sciences 12 3
[6] Turiman P, Omar J, Daud A M and Osman K 2012 J. Procedia-Social and Behavioral Sciences 59 110
[7] Husin W, Arsad, N, Othman O, Halim L, Rasul M S, Osman K and Iksan Z 2016 Proc. Asia-Pacific Forum on Science Learning and Teaching 17 1
[8] Partnership for 21 2011 21st century skills available at http://www.p21.org/our-work/p21-framework
[9] Azwar S 2007 Validitas dan reliabilitas (Yogyakarta: Pustaka Pelajar) p 44-61
[10] Sukardi H M 2013 Metode Penelitian Pendidikan Tindakan Kelas Implementasi dan Pengembangannya (Jakarta: Bumi Aksara) p 21
[11] Arifin Z 2012 Evaluasi Pembelajaran (Bandung: Rosdakarya) pp 31
[12] Borg WR, Gall MG 1989 Educational Research: An Introduction 5th ed. (New York: Longman)
[13] Fatimah AWN, Suryani N and Yaminah S 2018 Int. J. of Multicultural and Multireligious Understanding 5 2
[14] Redhana I and Sya’ban S 2014 Proce. SENARI 2 1
[15] Brookhart S M 2010 *How to assess higher-order thinking skills in your classroom Chap 1* (ASCD)

[16] Huijser H and Kek M Y C A 2015 *Int J. Higher Education Research & Development* **30** 3

[17] Sugiarti 2017 *Penerapan model problem based learning dengan reading infusion untuk meningkatkan keterampilan abad ke 21* (Bandung: UPI)

[18] Boss S, Larmer J, Mergendoller J R 2013 *PBL for 21st century success: Teaching critical thinking, collaboration, communication, and creativity* (Boston: Buck Institute for Education) p 23

[19] Gorman M 2018 available at http://www.21centuryedtech.wordpress.com

[20] Bie 2018 available at http://www.jetpost.com/eportfolio/pbl/driving_question.htm