Assessing 21st century skills of fourth-grade student in STEM learning

N Nurlenasari*, D A M Lidinillah, A Nugraha and G Hamdu
Universitas Pendidikan Indonesia, Kampus Tasikmalaya, Jl. Dadaha No. 18, Kota Tasikmalaya 46115, Indonesia

*nurlenasari@student.upi.edu

Abstract. This Performance assessment instruments to measure 21st century competencies, namely 4C (critical thinking, creativity, collaboration, and communication). Critical thinking, creativity, collaboration, and communication competence can be integrated in STEM learning from an early age. The purpose of this study is to develop a 4C performance assessment instrument that can be used for measuring tools in STEM-based learning in elementary schools. The method that has been used is Educational Design Research (EDR) according to Mc Kenney and Reaves through three stages namely Analysis and Exploration, Design and Construction, and Evaluation and Reflection. This study included 11 observers who observed 5 students as samples. Data collection was conducted by means of interviews, expert judgment and questionnaire entries. The results of the study are instruments of student performance assessment. This instrument is used as an assessment tool that can be used to assess student performance that includes criteria, achievement scores, and descriptions of achievement scores. This development of performance appraisal instruments can be used as a form of assessment because most of the results show Asymp. Sig <0.05 which is categorized as alignment. This 4C performance assessment instrument is expected to be used in general in conducting performance assessments that measure 21st century competencies in STEM learning.

1. Introduction
The elementary school curriculum is currently aimed at developing 21st century skills, such as critical thinking skills, creative thinking skills, the ability to collaborate and communicate, (4Cs skills) and the ability to competence of information and communication technology media. The current Indonesia curriculum places more emphasis on developing student’s character, literacy, higher order thinking skills (HOTS), and 4C. The teacher needs to design a learning model that is in accordance with the characteristics of the curriculum at this time which in its implementation uses a thematic-integrative learning approach.

STEM (Science, Technology, Engineering, Mathematics) is actually an integration of the four disciplines thus producing a “meta disciplines”. Integrated STEM education refers to a new name for the traditional approach to teaching science and mathematics [1]. Thus, learning approaches STEM is very suitable to be applied in the curriculum at this time, because of STEM education is based on the concept of educating students in four areas by integrating and applying it in the real world context. STEM education can be introduced to elementary school students. The importance of STEM education at the elementary level is getting higher, although education teacher preparation programs do not focus
on STEM education [2]. In the context of primary and secondary education, STEM education aims to develop students who have the STEM literacy (STEM literate), the details are as follows. First, have knowledge, attitudes, and skills to identify questions and problems in real situations, explain natural phenomena and design phenomena, and draw conclusions based on evidence regarding STEM-related issues. Second, understand the characteristics of the STEM fields as forms of human knowledge, as well as the investigation and design that can be human. Third, have awareness of how the STEM fields in shaping the environment of material, intellectual and cultural. Fourth, has the desire to be involved in STEM-related issues and to act as a constructive, caring and reflective citizen by using ideas in science, technology, engineering and mathematics [3].

In addition to the learning approach, assessment is also an important component in learning. Assessment becomes an important thing because the assessment can be achieved determine whether or not a learning objectives have been formulated previously. Assessment is a series of activities carried out by the teacher to obtain and process information about the process and learning outcomes of students in a learning. This is stated more explicitly by Hamdu which states that assessment is a series of activities to obtain, analyze, and interpret data about the process and learning outcomes of students carried out systematically and continuously so that they become information and meaningful participation in decision-making [4]. Definition of assessment by Angelo and Croos states that assessment is a process that is designed to help teachers discover what has been learned by the students in the classroom and how the level of success they learn it [5]. One of the assessments that are suitable for use in STEM learning is 4Cs assessment, because 4Cs has become part of the current elementary curriculum which is expected to be able to improve the quality of education so that graduates are ready to compete in the global world. The description of the 4C is as follows with indicators that can be trained or developed as critical thinking skills needed in the 21st century that adapted from the 2013 EdLeader21 [6]:

Table 1. Indicator of 4c skills.

| Performance aspects                  | Indicator                                      |
|--------------------------------------|------------------------------------------------|
| 1. Critical thinking                | 1) Information and discovery                   |
| Critical thinking and problem solving (critical thinking and problem solving): the desire to find out through a process of analyzing systems thinking and evaluating a situation to make decisions through ideas, evidence, reasons, and information in an effort to solve problems. The development of critical thinking is done by giving a reason effectively, using systems thinking, make judgments and decisions, and solve problems. | 2) Interpretation and analysis                  |
|                                      | 3) Reasoning                                   |
|                                      | 4) Problem solving                             |
|                                      | 5) Constructing arguments                      |
|                                      | 6) Reflection                                  |
| 2. Creativity                       | 1) Idea generation                             |
| Creativity and innovation (creativity and innovation): fluency and flexibility in thinking and expressing thoughts, as well as the ability to modify (elaboration) or create something new (originality) in the form of the idea and the real work. Innovation is a new discovery through application, synthesis, reinterpretation, in the form of real ideas and works of creativity and innovation can be characterized by creative thinking, creative work, and innovation. | 2) Ide design and refinement                    |
|                                      | 3) Openness and courage to explore             |
|                                      | 4) Work creatively with others                 |
|                                      | 5) Creative production and innovation          |
|                                      | 6) Reflection                                  |
| 3. Communication                    | 1) Engaging in conversation and discussions    |
| Communication and collaboration: the ability to absorb, deliver, and connect with the information and ideas in a variety of modes of language (spoken, written, signed, and visual). | 2) Using 21st century communication tools       |
|                                      | 3) Listening                                   |
|                                      | 4) Communicating in diverse environments       |
|                                      | 5) Delivering oral presentations               |
|                                      | 6) Reflection                                  |
Table 1. Cont.

| 4. Collaboration | 1) Leadership and initiative |
|-------------------|-------------------------------|
| Collaboration: the ability to work in teams to achieve common goals, including the ability to build partnerships and consensus, and in preventing and managing conflicts. | 2) Cooperation |
| | 3) Flexibility |
| | 4) Responsibility and productivity |
| | 5) Use of tech tools for synchronous and asynchronous collaboration |
| | 6) Responsiveness |
| | 7) Reflection |

Tabel 1 shows that, preliminary study was conducted at an elementary school in Indonesia which implemented the current curriculum, obtained information about the application of performance appraisal at the elementary school, it can be concluded that the teacher already knows and applies performance appraisal but in the implementation there are still many difficulties experienced by teachers in carrying out performance appraisal due to incomprehension regarding performance appraisal. Teachers simply use performance assessment rubric that already exist in the teacher's guide without being able to develop it themselves. The performance assessment is subjective because the teacher only looks at the cognitive, does not see the skills students have. Therefore, research development 4C performance assessment in STEM learning to measure all the skills of the students.

2. Methods

This research is part of a research team consisting of design, media, and assessments that are interrelated with each other. This article will discuss specifically the assessment used in the learning. this study used the Educational Design Research method. Design research is defined as the systematic analysis, design and evaluation of educational interventions with the dual aim of generating research-based solutions for complex problems in educational practice, and advancing our knowledge about the characteristics of these interventions and the processes of designing and developing them [7].

![Figure 1. Generic model of EDR.](image)

Figure 1 shows that the process of developing learning devices in the form of performance appraisal instruments refers to the EDR development model by Mc Kenney and Reeves which has three steps, namely: analysis and exploration, design and construction and evaluation and reflection [8].

The subjects of this study were elementary school students in the city of Tasikmalaya with a purposive sampling technique of 50 people. Data collection techniques in this study were interviews, expert judgment, and questionnaires. The instruments used in this study were interview guidelines, expert validation sheets, and questionnaires for observers. Interviews were conducted to grade IV
teachers, expert judgment was carried out when validating the feasibility of the product before testing, and a questionnaire was given to teachers and students who acted as observers after testing the product.

In this study, the validity test of the research instrument also consisted of an analysis of expert validation sheets and an analysis of the alignment (concordance) of Kendall. Validity test was used to test and determine the validity of an instrument. After consulting with experts, it is then tested, and analyzed by item analysis. Item analysis is done by calculating alignment with kendall test. Kendall test is used to determine the extent to which two sets of ranks of n individuals are aligned or not. This kendall test is carried out using the SPSS program (Statistical Package for Social Sciences) version 16.0. Kendall test has a hypothesis where \( H_0 \) means that there is no agreement or harmony between the observers in assessing attributes, and \( H_1 \) means that there is agreement or harmony between the observers in assessing attributes. the basis for making decisions is as follows: If the probability is > 0.05 then \( H_0 \) is accepted or if the probability is < 0.05 then \( H_0 \) is rejected.

3. Results and discussion

3.1. Analysis and exploration

The results of the analysis and exploration conducted by researchers on the learning process in the classroom it can be said that meeting the 2013 curriculum requires the development of 21st century (4Cs), and the assessment used must be an assessment that can measure the whole these skills. But the fact in the field found that the assessment carried out still tended to stick to student learning outcomes (test questions).

Based on the results of the literature study, researchers can conclude that the performance appraisal carried out by the teacher is still based on the cognitive domain only, whereas in the curriculum 2013 the assessment used is authentic assessment where the assessment pays attention to the balance between competency attitudes, knowledge and skills.

3.2. Design and construction

3.2.1. Design of learning design. In this step the researcher designed a 4C performance assessment instrument developed in STEM learning. Phase 4C developing an assessment instrument performance is focused on several activities, namely: create a design principle, making the grating, and designing the initial prototype 4C performance assessment instruments. The researcher compiled a performance assessment instrument in the form of a 4C performance assessment rubric. The format used in the rubric adapted from "EdLeader21” in 2013 [6]. Where each aspect of performance was in a different (separate) rubric. Indicator aspects used are only three of each aspect of performance, because they are adjusted to the needs of the learning. Rubric assessment of individual performance for each aspect 4C consists of two columns, the columns of unity are the indicators aspects to be observed, the second column shows the score and a description of each score will be given to students, while for information places, situations and tasks that should be done by students are not included in the column, but are listed below the table. In the critical thinking rubric, the reasoning aspect indicator is not obtained from the observation results but from the test questions. To benchmark assess the test item using a rubric that consists of two columns of the column scores and descriptions.

3.2.2. Validation of learning design products. This expert validation is done by presenting two experienced educational evaluation experts to assess the product design design of the 4C performance assessment rubric. The expert who becomes the validator evaluates and provides suggestions and improvements to improve the product so that the product is logically valid. Based on the results of the validation of the two validators, the results showed that overall the development of this 4C performance assessment instrument can already be used but there are some suggestions including rubrics adapted to the learning steps in the RPP, adjusting the source from EdLeader21 not to adapt from the source, for
description each score is specified again, the description places, situations and tasks should be placed in the column next to the section to make it easier for users to do the assessment.

3.2.3. **Product revision learning design.** After validation of product performance assessment instrument validator 4C by experts, there have been improvements to the design of performance assessment instrument 4C. The product is repaired in accordance with the suggested improvements provided by the expert validator. The performance assessment rubric which initially consists of two columns, namely the indicator column and the score column along with the description, changes into three columns according to the suggestion from the validator. One additional column is the time statement column.

3.3. **Evaluation and reflection**

3.3.1. **The first phase of implementation.** The first phase of the trial was carried out on eight UPI students at Tasikmalaya Campus and three teachers from Sukamulya Elementary School. At the trial's first phase, eight students and three teachers to act as an observer. The first phase of the trial was conducted on Wednesday, July 18, 2018 for observers who attended the learning process directly at Sukamulya Elementary School. Furthermore, on Thursday, July 19, 2018 for observers who observe via video in the space of 12 UPI Kampus Tasikmalaya. In Rubik's performance assessment of individual 4C, because of the limitations of student performance appraisal then there are only five students were observed by the observer and united in the same group to facilitate the observers in conducting the assessment. The data from the score recapitulation given by 11 observers to 5 students observed can be known as follows:

| Performance aspects | Results |
|---------------------|---------|
| S1  | S2  | S3  | S4  | S5  |
| C1  | 0.054 | 0.001 | 0.097 | 0.001 | 0.050 |
| C2  | 0.002 | 0.002 | 0.005 | 0.057 | 0.006 |
| C3  | 0.016 | 0.001 | 0.050 | 0.069 | 0.001 |
| C4  | 0.039 | 0.009 | 0.055 | 0.108 | 0.054 |

3.3.2. **The first phase reflection.** Based on the table 2 trial results, the alignment of the test results (concordance) kendall we can conclude that the actual performance assessment instrument 4C is quite feasible to use, but still have some improvements to make the product can be optimized further. Aside from the alignment test, the researcher also reflected on the product through a questionnaire intended for observers. From the questionnaire there were several inputs for researchers, including those related to the use of more simplified words or sentences, such as the indicator Delivering an Oral Presentation to Delivering an Oral Presentation. Then in determining the indicators of performance aspects it is better to avoid the word "and" as an example of the indicator of Interpretation and Analysis, you should choose only one. Furthermore, for the description of each performance aspect indicator, it is further clarified to make a difference in each score. After finding out the shortcomings, the researchers then made improvements to the product shortages and tested it back to elementary school. After the repair process is completed, then the researchers re-tested performance assessment instruments 4C.

3.3.3. **The second phase of implementation.** Test 2 was carried out after the improvement or revision II of the product developed by the researcher, namely the 4C performance assessment instrument. The second phase of the trial was also carried out on eight students of the UPI Tasikmalaya Campus and three teachers from Sukamulya Elementary School who acted as observers. The second phase of the trial was conducted on Tuesday, July 24, 2018 for observers who attended the learning process directly at Sukamulya Elementary School. Then on Wednesday, July 25, 2018 for observers who observed through video in room 12 of the UPI Tasikmalaya Campus. The observers filled out the 4C performance...
assessment rubric that had been improved by the researchers. The second stage trial was conducted in a different class from the first phase of the trial, namely class IV B Sukamulya Elementary School. In this second stage, the testing of the individual 4C performance assessment rubric is still done by observing only five students as a sample put together in the same group to make it easier for the observer to carry out the assessment. The data from the score recapitulation given by 11 observers to 5 students observed can be known as follows:

| Performance aspects | Results |
|---------------------|---------|
| S1                  | 0.001   |
| S2                  | 0.000   |
| S3                  | 0.012   |
| S4                  | 0.000   |
| S5                  | 0.052   |
| C1                  | 0.174   |
| C2                  | 0.018   |
| C3                  | 0.018   |
| C4                  | 0.018   |
| C5                  | 0.018   |

Based on table 2 that have been carried out, from the results of the kendall alignment test it can be concluded that the 4C performance assessment instrument shows improvement from the first phase of the trial with the second phase of testing, and can be said to be feasible because with a sample of 5 probability value <0.05 so that it can be said that there is agreement or harmony between observers in assessing attributes. The 1 sample with a probability value of> 0.05 can be caused from several factors including a video that is less supportive to make it easier for the observer to conduct an assessment, the observer's lack of objectivity in assessing, and from the students themselves who tend to be passive so it is very difficult to observe. From the results of the questionnaire intended for observers it also showed good results. Overall, this 4C performance appraisal rubric can be used as an assessment tool to develop children's potential to fulfill 21st century skills. Then the description of aspect indicators is even simpler so that it is easier for users to understand.

After two trials and two revisions to the 4C performance assessment instrument, the 4C performance assessment instrument for STEM-based learning can be used for grade IV in elementary school. So that learning can be done optimally, developing 4C skills in accordance with the demands of 21st century education.

4. Conclusion
The use of performance appraisal in class IV in Indonesia is still rarely done, especially for the evaluation of individual 4C performance. The performance assessment is usually only done when students learn in groups only. The performance assessment carried out by the teacher is still subjective because in assessing the performance of students the teacher still sees from the cognitive domain without considering the skills students have. 4C performance assessment instruments have been obtained for STEM-based learning. The 4C performance appraisal instruments developed in the form of assessment rubrics are used as a reference in conducting performance assessments. The developed rubric is an individual 4C performance assessment rubric. The aspects of individual performance that are developed include critical thinking skills, creativity, communication skills, and the ability to collaborate. The scale used in the 4C performance appraisal rubric is rating scale with 4 score choices which are 4, 3, 2, and 1. Scores for students who show the highest performance are 4 and for students who do not show performance are 1. Validation of 4C performance appraisal instruments is done through expert validation by expert lecturers and kendall alignment tests. The results obtained from expert validation state that this 4C performance assessment instrument is worthy of being used as an assessment tool. Statistically the results of testing using SPSS version 16.0 indicate that the 4C performance appraisal instrument developed has alignment with its attributes. Overall the results of the Kendall concordance test are categorized in harmony because they show a probability value of <0.05. After making improvements based on the results of product trials, the final product is obtained in the form of a 4C performance assessment rubric for STEM-based learning.
Acknowledgement
I would like to express my special thanks of gratitude to my lecturer (Prof. Nuryani Rustaman and Dr. Ahmad Mudzakir) who helped me in doing this research. I would also like to thank my parents, husband and friends who help me a lot in finalizing this research.

References
[1]  E L Mann, R L Mann, M L Strutz, D Duncan, and S Y Yoon 2011 Integrating Engineering Into K-6 Curriculum J. Adv. Acad. 22(4) 639–658
[2]  E Smyrnova-trybulska, N Morze, P Kommers, W Zuziak, and M Gladun 2016 Educational Robots In Primary School Teachers’ And Students’ Opinion About St Em 197–204
[3]  R W Bybee 2013 The Case for Education Challenges and Opportunities
[4]  G Hamdu 2016 Perangkat Pembelajaran Berbasis Masalah Secara Tematik di Sekolah Dasar (Bandung: Pelangi Press)
[5]  Y Abidin 2016 Revitalisasi Penilaian Pembelajaran Dalam Konteks Pendidikan Multilaterasi Abad Ke-21 (Bandung: Refika Aditama)
[6]  EdLeader21 2013 EdLeader21 4Cs Rubrics Grades 3rd/4th
[7]  Tjeerd Plomp 2013 Educational Design Research: An Introduction in Educational Design Research : part a : An Intriduction T Plomp and N Nieveen Eds Enschede the Netherlands 16
[8]  S McKenney and T C Reeves 2014 Educational design research in Handbook of Research on Educational Communications and Technology: Fourth Edition 131–140