Development of Contextual Teaching and Learning-Based Test Instruments to Improve 21st Century Skills in Students

Atika Ulya Akmal¹*, Festiyed²

¹Elementary School Teacher Study Program FIP, Universitas Negeri Padang, Padang, Indonesia.
²Physical Education FMIPA, Universitas Negeri Padang, Padang, Indonesia.

Abstract: The purpose of this study was to develop a test instrument for the basic physics concept course which was compiled based on the CTL learning model to improve students' 21st century skills (4C). This type of research is R&D and the design used by ADDIE. The subjects of this study were 25 semester 1 PGSD UNP students who worked on 33 questions on the instrument. The data analysis technique used is quantitative analysis by calculating the validity, reliability of the questions, discriminating power and level of difficulty. The results of this study were 78% valid questions and 22% invalid questions. Reliability of 0.933 which is included in the high reliability category. It is known that 18% of the questions are in the difficult category, 61% are in the medium category and 21% are in the easy category. The power of difference test obtained 6% of the questions included in the very bad category, 3% of the questions included in the bad category, 27% in the sufficient category, 61% in the good category and 3% in the very good category.

Keywords: CTL Model; Test Instruments; The 21st Century Skills

Introduction

In the current 21st century, education is expected to be able to keep up with the times in the field of science which is very rapid. Since technology, science, and talents have advanced so much in the 21st century, many aspects of life-including information and lifestyle have changed (Haug & Mork, 2021; Oktrisma & Ratnawulan, 2021). The advancement of science and technology has also had an impact on Indonesia as several aspects of Indonesians' life have changed. In response (Mashudi, 2021) stated that it is essential to develop top-notch human resources, particularly for students. It is intended that through teaching students 21st century skills in lectures, every student would be prepared to live in the 21st century with all of the chances and problems that will come with the era of information and technological advancements. Students must also possess the so-called 4Cs, which stand for communication, cooperation, critical thinking, and creative thinking.

The existence of this 21st century era encourages students to have strong abilities and skills to face life's challenges in the future. There are 9 the 21st century skills that must be mastered by students to answer the challenges of the times, namely knowledge and character to direct the formation of students' mindsets and character. Second, self-mastering or mastery of self-control. Third, Intellectual impressive or thinking process skills, students must have intellectual thinking skills in all conditions and challenges. Fourth, the capacity for interpersonal and communication exchanges. Fifth, leadership. Sixth, foreign language proficiency or ability to master a foreign language, this era at least requires foreign language or English skills, and it is a must have for students. Seventh, entrepreneurship is in line with the vision. Eighth, strengthening the history of civilization or strengthening the history of civilization. Ninth, education must be understood as a process of humanizing humans, this must not be forgotten from the true nature of education (Mardhiyah et al., 2021).

Today's education is often carried out with digital learning that develops along with technological developments. This development touches on all subjects.
including the Elementary School Physics Basic Concepts course. Learning the Elementary Physics Basic Concepts course requires digital learning. It is supported by (Mallidis-Malessas, P., Iatraki, G., & Mikropoulos, 2021; Prabasari et al., 2021) who contend that supplemental digital learning enhances pedagogical structures in science education and stimulates interactive learning in physics. The use of this technology is one of the characteristics of learning to develop 21st century skills in students. However, the reality on the ground is that no special efforts have been made to acquire 21st century skills that can be taught to university level students (Arsanti et al., 2021). Observations of first-year students at Padang State University are consistent with this. Observations led to the discovery of several significant ideas. That is, a lack of projects in the Elementary School Physics Basic Concepts course prevented students' creativity and innovation from developing to their full potential, the use of tools in apprenticeship activities did not directly integrate with ICT and the amount of lab equipment available was out of proportion to the number of students, which decreased students' ability to interact with their surroundings and educational resources.

According to the findings of educator interviews, instructors hardly ever provide instruction to hone students' critical thinking, creativity, communication, and teamwork skills. Only in front of the class do teachers provide lesson material using the discussion and Question & Answer format. Students appear less engaged in the learning process on average because they ask questions or voice comments less frequently. According to the findings of student interviews, it is known that students are typically less interested in and enjoy basic physics concepts courses, because the subject matter is difficult to solve and because basic physics concepts courses contain many terms that are difficult to understand. Students receive learning through the lecture method.

To build students' 21st century skills, learning reform that puts students first instead of teachers first is the solution (Ramdani et al., 2019; Redhana, 2019). Understanding the significance of developing 21st century abilities is a crucial ability for educators to foster in their students' learning. Learning models can be used to create abilities for the twenty-first century (Resti, 2018; Wrahatnolo & Munoto, 2018). The Contextual Teaching and Learning (CTL) learning model is one that can be used to acquire 21st century abilities. The CTL learning model is anticipated to be able to help students acquire 21st century abilities like teamwork, communication, creativity, and critical thinking (Asrizal & Utami, 2021; Yunita & Aafa, 2020). The CTL model is also able to package student aptitude testing even more effectively because abilities will be sharpened by facilities for learning activity management and discovery of more tangible and applicable learning experiences for students (Dewi & Primayana, 2019; Kurniasari et al., 2018).

Moving on from previous research by (Welerubun et al., 2022) stated that his study's findings, which used the CTL model to examine 21st century abilities, demonstrated how critical thinking, creativity, communication, and teamwork grew as a result of using the CTL paradigm. Research by (Shantia & Lufri, 2021) it has been demonstrated that using the CTL approach can help students develop 21st century abilities, including the ability to think critical. Research The 21st century national education paradigm also emphasizes the importance of contextual learning abilities as a set of abilities that 21st century human resources must possess (Husain & Kaharu, 2020). By relating disciplines to their context in daily life, CTL can help students understand the significance of an academic lesson they are studying (Sidiq et al., 2021). The constructivism, questioning, inquiry, learning community, modeling, reflection, and authentic assessment are the seven key elements of CTL model (Sulistiani, 2020; Yani et al., 2021).

Students will acquire 21st century skills well if they use good evaluation tests. Efforts to develop a 21st century skills test instrument (4C) must meet the eligibility of a good test instrument. A good test is a test that has valid, reliable characteristics. A test can be said to be valid if the test can be accurately measured. The test is said to be reliable if it gives the right results when tested many times. A test is said to have high practicability if the test is practical, i.e. easy to implement, easy to examine, and equipped with clear instructions (Tersiana, 2018; Yuwandra & Arnawa, 2020).

Based on the description above, a test development research was conducted to improve students' 21st century skills. The purpose of this study was to develop a test instrument for the basic physics concept course which was compiled based on the Contextual Teaching and Learning (CTL) to improve students' 21st century skills (4C).

**Method**

This research is an R&D research with the ADDIE (Analysis, Design, Development, Implementation, and Evaluation) approach. This research was conducted at Padang State University on semester 1 PGSD students, totaling 25 students. This research is in the form of developing a test instrument that is arranged in such a way in the basic concepts of physics course in elementary school. The data analysis technique used in...
this study is quantitative analysis by calculating the validity, reliability of the questions, discriminating power, and difficulty level. This model uses 5 stages that is analysis stage, Design, Development, Implementation, and Evaluation. The following is the procedure for implementing the ADDIE development model.

![Figure 1. ADDIE development model](image)

Even though the development procedure is shortened, it already includes a testing and revision process to ensure that the product developed meets the product criteria. Testing is done by a team of experts, individual research subjects, limited and wide scale (field), and revisions are made to improve the final product, excellent, empirically tested, and no mistakes. This study was to develop a test instrument for the basic physics concept course which was compiled based on the Contextual Teaching and Learning. To produce a good test instrument, test analysis tests are used where this analysis stage is included in the development stage of the ADDIE model. This development is used to view improvement in students' 21st century skills. This test instrument went through a validation process by the validator. The validator is a person who is an expert in a particular field who has the right to justify or validate an instrument. The validators used in this study were 3 expert validators who were competent in their fields. The results of the validation of the 3 validators are as follows:

<table>
<thead>
<tr>
<th>Table 1. Validation Results by the Validator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validators</td>
</tr>
<tr>
<td>Validators 1</td>
</tr>
<tr>
<td>Validators 2</td>
</tr>
<tr>
<td>Validators 3</td>
</tr>
<tr>
<td>Average</td>
</tr>
</tbody>
</table>

Based on the aforementioned validation results, it was determined that the average validator percentage was 79.66%, indicating that the test instrument's 33 questions can be used and will remain valid with very minor adjustments. Then the test is tested in the field or tested on students to recalculate its validity. The calculation results are as follows:

<table>
<thead>
<tr>
<th>Table 2. Question Item Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Question</td>
</tr>
<tr>
<td>1, 2, 4, 5, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 23, 25, 26, 28, 29, 30, 31, 33</td>
</tr>
<tr>
<td>r table = 0.413</td>
</tr>
<tr>
<td>3, 6, 9, 22, 24, 27 and 32</td>
</tr>
<tr>
<td>r table = 0.413</td>
</tr>
</tbody>
</table>

It may be seen from the table above that there are 7 invalid items and 26 legitimate questions. The results of the r count computation are compared to the r tables, which are produced using the phrases \( n - 2 \), where \( n \) is the total number of respondents. The outcomes of validating the items are shown in the diagram below:

![Figure 2. Results of validation of question items](image)

**Result and Discussion**

This research produced a product in the form of a multiple choice test instrument based on the CTL model on the basic concept of elementary school physics. There were 33 tests used. Several experiments were conducted to determine the viability of this test instrument, including validity, reliability, difficulty level, and differential power testing. The series of tests carried out are described as follows:

**Validity Test**

The test instrument used in this study was prepared with the guidance of a test instrument grid which of course was made based on indicators related to 21st century skills.
value in the high reliability value classification (0.81 < r1.1 ≤ 1.00) which indicates high reliability.

**Difficulty Level Test**

The quality or level of difficulty possessed by each of these items can be used to measure the quality of the test items. The ability of the pupils to react to these questions, rather than from the viewpoint of the educator as the question's author, determines the complexity of the question. There are three categories of difficulty level questions, namely difficult, medium and easy. From the calculation results, the difficulty level of the test is obtained as follows:

<table>
<thead>
<tr>
<th>Item Range</th>
<th>Value Range (p)</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>2, 4, 13, 15, 16, 21, 24</td>
<td>0.71 – 1.00</td>
<td>Easy</td>
</tr>
<tr>
<td>1, 5, 6, 7, 8, 9, 10, 11, 12, 14, 17, 18, 20, 23, 25, 26, 28, 29, 31, 33</td>
<td>0.31 – 0.70</td>
<td>Currently</td>
</tr>
<tr>
<td>3, 19, 22, 27, 30, 32</td>
<td>0.00 – 0.30</td>
<td>Hard</td>
</tr>
</tbody>
</table>

From the table above it is known that there are 7 items belonging to the Easy category, 20 items belonging to the medium category and 6 items belonging to the difficult category. The following is a diagram of the results of the difficulty level of the items:

**Power Difference Test**

Discriminating power aims to assess the ability of an item to separate respondents who are considered capable from respondents who are considered less capable or less accomplished. There are five different power categories, namely very bad, bad, enough, good, very good. From the calculation results, the differential power of the test is obtained as follows:

<table>
<thead>
<tr>
<th>Number of Question</th>
<th>Grade Range (D)</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>9, 27</td>
<td>D: &lt; 0.00</td>
<td>Ugly as hell</td>
</tr>
<tr>
<td>32</td>
<td>D: 0.00 – 0.20</td>
<td>Bad</td>
</tr>
<tr>
<td>2, 5, 6, 16, 22, 24, 25, 26, 33</td>
<td>D: 0.21 – 0.40</td>
<td>Enough</td>
</tr>
<tr>
<td>1, 3, 4, 7, 8, 10, 11, 12, 13, 14, 15, 18, 19, 20, 21, 23, 28, 29, 30, 31</td>
<td>D: 0.41 – 0.70</td>
<td>Good</td>
</tr>
<tr>
<td>17</td>
<td>D: 0.71 – 1.00</td>
<td>Very well</td>
</tr>
</tbody>
</table>

From the table above it is known that there are 2 items in the very bad category, 1 item in the bad category, 9 items in the moderate category, 20 items in the good category and 1 item in the very good category. The following is a diagram of the different power results of the questions:

The development of this test instrument is needed to determine the extent to which the test can be used and measure the development of the instrument. The test instrument is said to be suitable for use when it has been tested for validity, reliability, level of difficulty and differentiability of each item on the test. The researcher validated the instrument with the expert validator and obtained an average validation value of 79.66%, which means that the test can be used with minor improvements. Researchers designed a matter of 33 items. It turned out that after the validity test was carried out in the field, there were 7 questions that were invalid and 26 questions that were valid. By contrasting the value of the r table and the r count of each item, it is possible to determine if the query is valid or not. There are 78% legitimate queries and 22% false ones.

The reliability test was carried out to see whether the question was reliable or not. The calculations that have been done show that the questions in this study are included in high reliability with a value of 0.933. The different power level test was carried out to find out which category each item belongs to. From the calculation results, there are 18% questions in the difficult category, 61% questions in the medium category and 21% questions in the easy category. It is supported by (Fatimah, 2019) which says that the consistency of a set of questions can be used to determine their level of complexity, and the compilation of respondents' responses can indicate if a group of questions is easy, medium, or challenging in terms of difficulty.

The power of difference test was carried out to see whether each item had a significant similarity or not (Handayani & Iba, 2020). From the calculation of the differential power test performed, it was found that 6% of the questions were in the very bad category, 3% of the questions were in the bad category, 27% were in the sufficient category, 61% were in the good category and
3% were in the very good category. It is supported by (Iskandar & Rizal, 2018) which says that the greater an item's coefficient of discriminating power, the better it is able to distinguish between respondents who have mastered competence and those who have not.

Based on the overall results of the research, 26 questions were produced that met the eligibility criteria for a good instrument covering aspects of validity, reliability, discrimination, level of difficulty so that they were suitable for training and measuring 21st century skills of PGSD UNP students in the basic concepts of elementary school physics course.

Conclusion

The development of test instruments carried out in this study concluded that the test instruments were appropriate or good to use as instruments that could improve 21st century skills in students. This test instrument has fulfilled the test requirements of good validity, reliability, level of difficulty and differential power of questions. There are 26 questions that meet the predetermined requirements so that they become proper and good instruments. The Contextual Teaching and Learning (CTL) learning paradigm is integrated into the design of this test instrument, and the HOTS question guidelines are also used to pick the proper operative verbs. This produces a good test instrument that can be applied later.

Acknowledgments

Thanks to Padang State University and the elementary school teacher education department for facilitating this research. Thank you also to all parties who assisted in completing the research, students and lecturers who were involved in the research process.

Author Contributions

Conceptualization, Atika Ulya Akmal and Festiyed; methodology, Festiyed; validation, Ahmad Fauzan and Abdul Razak; formal analysis, Atika Ulya Akmal and Festiyed; investigation, Atika Ulya Akmal and Festiyed; resources, Atika Ulya Akmal; data curation, Atika Ulya Akmal; writing—original draft preparation, Atika Ulya Akmal and Festiyed; writing—review and editing, Atika Ulya Akmal and Festiyed; visualization, Atika Ulya Akmal and Festiyed; supervision, Festiyed; project administration, Atika Ulya Akmal. All authors have read and agreed to the published version of the manuscript.

Funding

This research was not funded by any party, but using self-financing.

Conflicts of Interest

The authors declare no conflict of interest.

References


