Journal for the Education of Gifted Young Scientists, 7(3), 747-775, September 2019
e-ISSN: 2149- 360X
http://jegys.org

Research Article

Development of Physics-Tier Tests (PysTT) to Measure Students' Conceptual Understanding and Creative Thinking Skills: A Qualitative Synthesis

Himawan PUTRANTA[1] & SUPAHAR[2]

Received: 4 July 2019 **Accepted:** 14 September 2019

Abstract

This paper is based on the background of the problem of the low high order thinking skills in students, especially in the skills to think creatively and conceptual understanding. Conceptual understanding that students have in relation to physics learning material has an important role in developing students' high order thinking skills in solving problems of daily life creatively. The method used in this research is descriptive qualitative research method with literature studies. The results of this study were obtained a synthesis of physics-tier tests (PysTT) to measure conceptual understanding and students' creative thinking skills, which are the basis for the development of PysTT based on real-life problems experienced by students realistically. In addition, PysTT is also one of the assessment instruments whose development is based on aspects of students' conceptual understanding of physics matter in everyday problems by prioritizing aspects of identifying and formulating problems, identifying scientific evidence and phenomena, arranging conclusions, and communicating conclusions creative. In addition, the results of the validity and reliability physics two-tier tests to measure students' conceptual understanding and creative thinking skills were 0.81 and 98.82%, which were included in the very good and reliable category. The ability to conceptual understanding and creative thinking skills of students who are measured using the physics two-tier test is very good because it is more than the value of the specified minimum completeness criteria.

Keywords:

physics-tier tests, conceptual understanding, creative thinking skills

To cite this article:

Putranta, H., & Supahar. (2019). Development of Physics-Tier Tests (PysTT) to Measure Students' Conceptual Understanding and Creative Thinking Skills: A Qualitative Synthesis. *Journal for the Education of Gifted Young Scientists, 7*(3), 647-775. DOI: http://dx.doi.org/10.17478/jegys.587203

[1] M.Pd. Candidate, Department of Physics Education, Graduate School of Yogyakarta State University, Yogyakarta, Indonesia, E-mail: himawanputranta.2017@student.uny.ac.id /ORCID ID No: 0000-0003-2646-8237
[2] M.Si., Dr., Department of Physics Education, Graduate School of Yogyakarta State University, Yogyakarta, Indonesia, E-mail: supahar@uny.ac.id /ORCID ID No: 0000-0002-2486-5549

Introduction

One aspect of a country can be said to be advanced, if the aspect of education has become a major need for the people of that country (Porter and Kramer, 2018). Education can not only be implemented in schools, but can also be carried out in the family environment, as well as educational institutions. Meanwhile, the goal of education universally is to educate all human life. The purpose of education becomes an encouragement for the community to have a more competitive quality of human resources (Gruenewald and Smith, 2014). One effort that can be done in improving the quality of human resources is more competitive by arranging learning objectives and appropriate evaluation instruments.

In learning physics in high school, students are required to achieve maximum learning goals in terms of developing the skills to conceptual understanding and think creatively by studying every problem that occurs in life using appropriate physics concepts (Collins, 2014; Barrow, 2015). Meanwhile, with the right learning objectives tailored to the problems that occur in life, it needs development in terms of evaluating the skills of students. Evaluation instruments used to measure each of the skills possessed by one of them by using a test (Newcombe and Shipley, 2014; Loewenthal and Lewis, 2018). Evaluation instruments such as test instruments should be developed with careful planning to measure conceptual understanding and high order thinking skills, i.e creative thinking (Aizikovitsh-Udi and Cheng, 2015). Therefore, with the aim of physics learning and evaluation instruments in the form of appropriate tests, it is expected that students can optimize conceptual understanding and thinking skills to solve physics problems creatively.

Regarding creative thinking skills, conceptual understanding is the most basic thing students must have in learning physics. This is because conceptual understanding is the skills to understand physics concepts appropriately/not misconception, in the sense of understanding the concept of physics that is universally applicable throughout the world (Kurniawati et al., 2017; Mądrala et al., 2017). If students have understood the concept of physics appropriately, it is possible for students to solve various problems in life using the physics concepts they have understood. After students are able to solve various problems in life using physics concepts continuously, they can develop and modify more diverse ways of solving using their more creative mindsets or strategies (Furberg, 2016; Sadiqin et al., 2017). Creative thinking skills is basically a skills that arises because often in solving problems using different ways (Kiryak and Çalik, 2017). Therefore, one effort that needs to be done to improve the creative thinking skills is too often practice in solving a physics problem in life.

Creative thinking skills are part of high-order thinking skills (HOTS) proposed by Bloom in addition to the low-order thinking skills that students must possess

(Istiyono et al., 2014). However, based on the TIMSS results show that the average achievement of Indonesian students' physics learning outcomes in terms of cognitive aspects (knowing, applying, reasoning) is still low (Efendi, 2010). In addition, the TIMSS results also show that the tendency of achievement of Indonesian students' physics learning outcomes always decreases in every cognitive aspect so that Indonesian students' physics skills must be improved in all aspects, especially on aspects of reasoning by equipping students with conceptual understanding and creative thinking (TIMSS & PIRLS International Study Center, 2012; Istiyono et al., 2014). In line with these findings, Luangrath et al. (2011) found that the results of the assessment of conceptual understanding and creative thinking of students towards Mechanics were still low.

Therefore, in general, the achievement of Indonesian secondary school students' physics learning outcomes in the international arena which requires the conceptual understanding and creative thinking skills is low (Wibowo and Suhandi, 2013). Achievement of low physics learning outcomes can be caused by physics learning activities or inappropriate test instruments (Rahmatan et al., 2012). In this case only the test instruments will be discussed, because the right test instrument can generate students to learn by creative thinking skills based on their conceptual understanding of physics that are in everyday problems appropriately. Furthermore, assessment of all aspects of skills possessed by students can be done in two methods, i.e verbally or in writing (Apino and Retnawati, 2017). Written assessment is done by using a test instrument that is done in writing in the form of choosing an answer and filling in the answer. Meanwhile, written test questions whose answers are done by choosing answers, i.e multiple choice, two choices (right-wrong, yes-no), matchmaking, and the causation. Meanwhile, it should be noted that the assessment model also influences creative thinking skills and conceptual understanding (Van den Berg, 2008). Meanwhile, the fact that multiple choice tests are more widely used to measure students' conceptual understanding and creative thinking skills than other forms of testing (Istiyono et al., 2014).

Furthermore, nowadays multiple choice tests developed to measure students' conceptual understanding and creative thinking skills do not only require to choose one correct answer from several answer choices provided. However, it consists of a multiple choice-tier test containing at least two choices, i.e the choice of answers and the choice of reasons for the answer (Wilcox et al., 2015). This aims to encourage students to express/choose reasons from the answers they choose for physics problems (Barniol and Zavala, 2016). In other words so that students are able to provide creative reasons for the answers they choose based on the physics concepts they understand. Meanwhile, the reason for the development and modification of multiple choice tests became multiple choice-tier tests on physics subjects, i.e the physics material tested could cover most of the physics learning

materials, choice of answers and choice of reasons students could be corrected easily and quickly, answers to each question it is certainly true or false, so the assessment is more objective (Sudjana, 1990; Caballero et al., 2017). However, multiple choice-tier tests also have weaknesses as well as multiple choice tests in general, i.e allowing students to guess answers even though not as large as in multiple choice questions in general and students' creative thinking processes cannot be seen clearly (Sudjana, 1990; Caballero et al., 2017).

Furthermore, in multiple choice-tier tests there is also cheating, for example students are still collaborating with other students during the test, then the format for either numbering or test sequences performed by each student should be different (Ding, 2014). Meanwhile, in assessing the results of the work of students in working on multiple choice-tier tests it is necessary to be based on assessment rubrics that assess each stage that students can complete (Bates et al., 2014). Just as in multiple choice-tier test assessment consisting of choices of answers and reasons, students who get the highest score are of course obtained when students are able to choose the answers and reasons in the item correctly.

Based on the description, then to measure the conceptual understanding and creative thinking skills in physics subjects are used multiple choice-tier tests, such as second-tier multiple choice tests that require to answer/choose the right answers and reasons. For this reason, it is necessary to develop an assessment instrument for conceptual understanding and creative thinking skills consisting of test instruments and assessment guidelines/assessment rubrics (Docktor et al., 2016). However, the most basic stage before making an instrument for assessing the conceptual understanding and creative thinking skills is to conduct a literature review of the results of other people's research or studies of appropriate books. Therefore, it is obtained a synthesis or new author's view of tier test instruments which aims to measure the conceptual understanding and students' creative thinking skills towards physics problems. Thus, there is a need for descriptive qualitative research that contains a detailed synthesis of the characteristics of tier test instruments aimed at measuring students' conceptual understanding and creative thinking skills. The existence of this synthesis is expected to facilitate the development of tier test items and physics assessment guidelines in future studies.

Method

In accordance with the objectives in this study to obtain new ideas or new syntheses that form the basis for the development of PysTT instruments used to measure students' conceptual understanding and creative thinking skills, the research method used is a qualitative research method as a basis for discussing research results. Qualitative research method itself is a research method that analyzes data in the form of information, translations in oral or written form which are then

connected with other data to get clarity about the truth, so that new ideas or synthesis of a particular concept are obtained (Neuman, 2011; Grbich, 2013) Furthermore, the research approach used is a descriptive approach, namely the procedure of problem solving by describing the state of the research subject both in written and oral form based on facts or research results (Flick, 2013; Grbich, 2013). In this study which is a qualitative descriptive study, the data analyzed in the form of data in the form of words or writings originating from books or the results of research by experts who are competent in their fields, especially in the field of physics-tier test instruments, understanding of physics concepts, and creative thinking skills.

Meanwhile, the sample used as the subject of this study were primary sources from books and secondary sources from the results of the research, each of which was published after 2010. The sample used in this study was not only books, but also used scientific articles as a result of the study because the two samples were triangulated to strengthen the new synthesis obtained by the researcher. The sample details used in this study are, on aspects of conceptual understanding of students using a sample of 7 copies and scientific articles from 7 relevant studies, aspects of students' creative thinking skills using 6 sample books and 7 scientific research articles. Relevant research, physics-tier test instruments (PysTT) using scientific article samples from the research results of 8 relevant studies, aspects of physics concept understanding tests using book samples of 3 copies and scientific articles from 5 relevant studies, and aspects of thinking skills tests creative students in physics subjects use a sample of 3 copies of the book and scientific articles from the research as many as 5 relevant studies.

Meanwhile, the technique of determining samples in the form of books and scientific articles that are relevant research results as the subject of this study using a purposive sampling method. As stated by Marshall and Rossman (2014) that the purposive sampling method is one method or technique used in taking research samples/research subjects that are not based on levels and carried out randomly, but based on the existence of certain objectives. The book which is used as the primary source or the main sample in this study is chosen based on the content of the book, the number of quotes, and the suitability of the purpose of this study. Meanwhile, the scientific articles selected from the research were grouped based on novelty, innovation, the number of quotes, and the suitability of the purpose of this study. Meanwhile, data collection techniques used in this study use literature studies by selecting and studying research samples in the form of books and scientific articles that are relevant research results in accordance with the considerations and provisions described earlier.

After the data from the sample in the form of books and scientific articles of relevant research results are analyzed, the next step is to conclude the data

specifically so that new ideas or new syntheses are obtained in accordance with the research objectives, namely the synthesis of PysTT instruments used to measure students' conceptual understanding and creative thinking skills. Meanwhile, the details of the method or steps taken in this study can be shown in Figure 1.

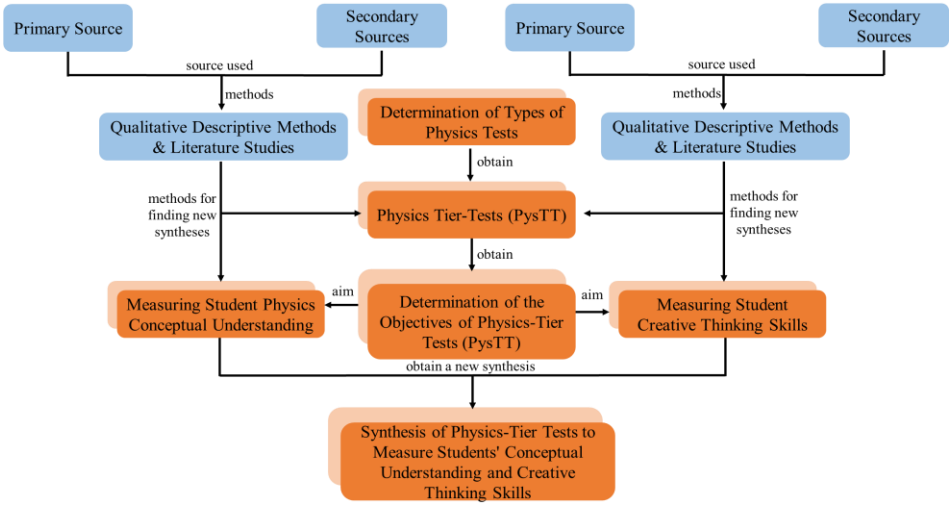


Figure 1.
Steps of the Qualitative Descriptive Methods and Literature Study

Research Sample

This research was conducted at SMA N 9 Yogyakarta. SMA N 9 Yogyakarta is a state senior high school in the Special Region of Yogyakarta, Indonesia. This research was conducted at SMA N 9 Yogyakarta in February-May 2019. In addition, the research subjects were 10th grade students of Mathematics and Natural Sciences at SMA N 9 Yogyakarta. MIPA itself is an abbreviation of mathematics and natural sciences, where the abbreviation is often used in Indonesia. In this study, the classes used amounted to two classes. The number of students in each class is different, in class 10 MIPA 1 there are 31 students and class X MIPA 2 has 30 students.

Meanwhile, the technique of determining the sample as the subject of this study uses a purposive sampling method. Marshall and Rossman (2014) state that the purposive sampling method is a technique or method of taking research subjects that are not based on level and random, but are based on the existence of certain objectives. Students used as research samples are selected based on the suitability of the physical material used in research or measurement, namely material momentum and impulses. Meanwhile, the data collection techniques used in this study use reasoned multiple choice questions or physics tier-two test questions that the researchers developed as can be shown in Figure 2.

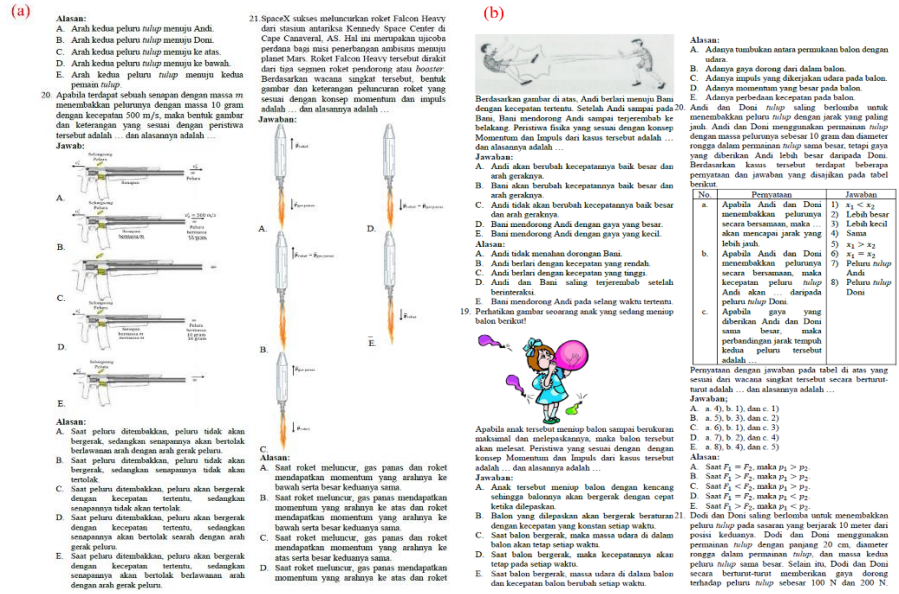


Figure 2.
Multiple Choice Test Items With Reason or Physics Tier-Two Tests, (a) to Measure Creative Thinking Skills, and (b) to Measure Conceptual Understanding

Meanwhile, research documentation when students work on reasoned multiple choice test questions or physics two-tier tests can be shown in Figure 3.

Figure 3.
Students Complete Reasoned Multiple-Choice Test Questions or Physics Two-Tier Tests

Instrument Analysis

Meanwhile, the reasoned multiple choice test questions or physics two-tier tests that have been developed by these researchers are then analyzed for their validity and reliability. The feasibility of a reasoned multiple choice test question or physics two-tier tests that has been developed is obtained from the validation score given by the expert and practitioner validator and the results of student responses. Therefore, to analyze the feasibility of a reasonable multiple choice test item or physics two-tier tests obtained from the validator and student assessment results the responses were carried out using the Aiken V equation as shown in Equation 1.

$$V = \sum \frac{s}{[n(c-1)]} = \sum \frac{r - l_o}{[n(c-1)]} \quad (1)$$

In this case, s is the judge to *n*, l_o is the lowest validity score, c is the highest validity score, and r is the number given by the assessor to *n*. Meanwhile, Azwar (2015) stated that the criterion of the validity score of physics two-tier tests obtained based on equation 1 is divided into 5 as shown in Table 1 below.

Table 1.
Criteria Score of Aiken's V validity

Validity Score	Category
$0.8 \leq V \leq 1.0$	Very Good
$0.6 \leq V \leq 0.8$	Good
$0.4 \leq V \leq 0.6$	Quite Good
$0.2 \leq V \leq 0.4$	Bad
$V \leq 0.2$	Very Bad

After validation and student responses to physics two-tier tests are analyzed using Aiken V Equations, the next step is to analyze the reliability scores of physics two-tier tests. The results of the reliability analysis of the physics two-tier tests were obtained using the agreement percentage (PA) analysis. The way to determine the reliability of physics two-tier tests is to correct and evaluate student work on physics two-tier tests by two assessors and then test the level of agreement using the percentage agreement equation (PA) as shown in Equation 2 (Borich, 1994).

$$PA = \left(1 - \frac{A - B}{A + B}\right) \times 100\% \quad (2)$$

In this case, the PA is the value of the percentage of agreement, A represents a higher total score of the assessor, and B represents a lower total score of the assessor. Based on the value of PA can be known level of approval physics two-tier tests, provided that the percentage agreement value of ≥ 75% and it can be stated that both assessors agree or reliable.

Data Analysis

In this research, it will only measure the ability to conceptual understanding and creative thinking skills of students in two classes from SMA N 9 Yogyakarta. Equations that are used to measure the achievement of the ability to conceptual understanding and creative thinking skills of students using equation 3 below.

$$\text{Value} = \textit{the score obtained by each item} \times 4 \quad (3)$$

Results and Discussion

In this section we will discuss a number of things about synthesis that were successfully compiled by the author based on several primary and secondary sources or based on relevant research related to the concept of PysTT, basic theories of conceptual understanding, basic theories of creative thinking skill, concepts of test physics conceptual understanding, concepts test physics creative thinking skills. The synthesis results that have been obtained by this author will then be used in the preparation of PysTT instruments used to measure students' conceptual understanding and creative thinking skills in high school physics subjects.

The Concept of Physics-Tier Tests (PysTT)

The evaluation section in learning is one of the final parts in learning activities that are very important to know or measure every ability that is in students. In measuring each student's abilities, of course, requires a technique and assessment instrument, either verbally or in writing and using a test or questionnaire. In accordance with what has been discussed in the background of the problem in the previous section, by considering the strengths and weaknesses of the written test instrument (multiple choice test and essay test), the physics-tier test (PysTT) is chosen which demands not only student answers but also conceptual reasons creatively from a physics problem in everyday life.

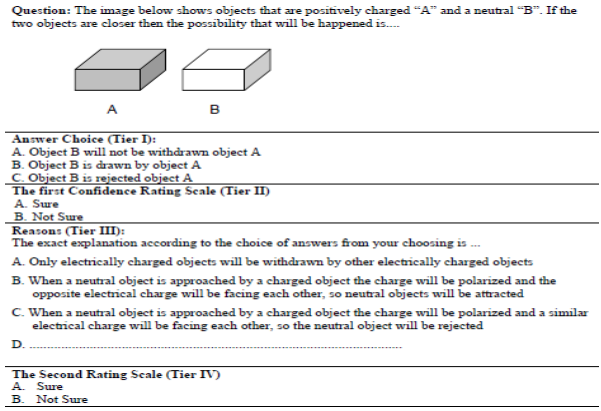
According to Winarti et al. (2017) states that physics assessment instruments in the form of two-tier tests are test instruments used to measure students' conceptual understanding of heat and temperature chapter in the form of reasoned multiple choice tests. This PysTT instrument consists of 20 items, in this case the first tier of each item consists of questions with four answer choices, and the second tier is the reason for the answer choice in the first tier that has been chosen. The example of a test instrument developed by Winarti can be shown in Figure 4.

1. **Mother put 100 grams of ice at 0 °C and 100 grams of water at 0 °C into a room with a temperature of 27 °C. After waiting long enough for the system to be balanced, which temperature is higher?**
   a. **Ice and water have the same temperature**
   b. **Water**
   c. **Ice**
   d. **Can not be determined**

Figure 4.
Example of a Test Instrument Developed by Winarti

In addition, this PysTT instrument is also effective in measuring students' conceptual understanding of heat and temperature chapter as evidenced by the Cronbach alpha reliability coefficient of 0.73. Furthermore, Hermita et al. (2017) states that assessment instruments used to measure conceptual understanding and creative thinking skills can also done by using a physics four-tier test assessment instrument. The four-tier test developed was adjusted to the concept of chapter static electricity with the first tier in the form of a choice of four answers, the second tier in the form

of a choice of four reasons, the third tier was the reason for the answer chosen, and the fourth tier was the reason for the reasons chosen. This assessment instrument in the form of a four-tier test is effective for assessing the conceptual understanding of the chapter on static electricity, as evidenced by the ability of the four-tier test instrument to distinguish between students who understand the concept of static electricity correctly and students who experience misconceptions in static electricity. The example of a test instrument developed by Hermita can be shown in Figure 5.



Question: The image below shows objects that are positively charged "A" and a neutral "B". If the two objects are closer then the possibility that will be happened is....

A B

Answer Choice (Tier I):
A. Object B will not be withdrawn object A
B. Object B is drawn by object A
C. Object B is rejected object A

The first Confidence Rating Scale (Tier II)
A. Sure
B. Not Sure

Reasons (Tier III):
The exact explanation according to the choice of answers from your choosing is ...

A. Only electrically charged objects will be withdrawn by other electrically charged objects

B. When a neutral object is approached by a charged object the charge will be polarized and the opposite electrical charge will be facing each other, so neutral objects will be attracted

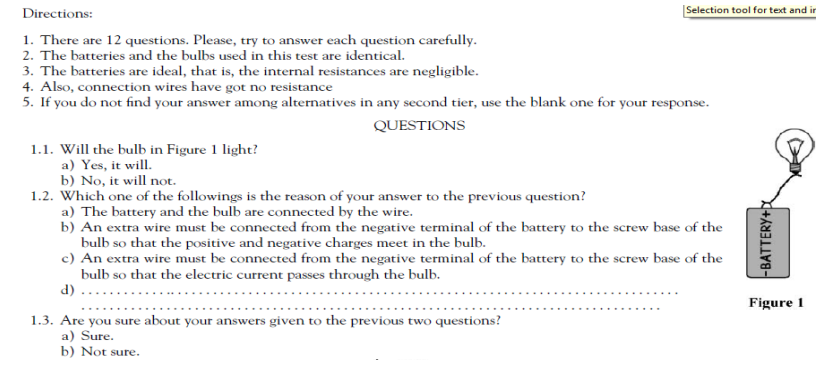
C. When a neutral object is approached by a charged object the charge will be polarized and a similar electrical charge will be facing each other, so the neutral object will be rejected

D. .....

The Second Rating Scale (Tier IV)
A. Sure
B. Not Sure

Figure 5.
Example of a Test Instrument Developed by Hermita

According to Peşman and Eryilmaz (2010) state that other PysTT instruments that can be used to measure students' conceptual understanding and creative thinking skills are using three-tier multiple choice tests. Physics assessment instruments are in the form of a three-tier test almost similar to a two-tier test, but there is one additional tier that asks students whether they are confident about their responses to the previous two tiers (Peşman and Eryilmaz, 2010; Yusrizal and Halim, 2017). This assessment instrument consists of 12 three-tier multiple choice questions that are useful for assessing students' conceptual understanding of simple electrical circuits. The first tier is a multiple choice question with 3-5 answer choices and the second tier presents several reasons for the answers selected at the first tier. Meanwhile, at the third tier in the form of choices about confidence in the answers that choose students on the previous two tiers. This physics assessment instrument in the form of a three-tier test is effective for assessing the students' conceptual understanding related to the simple electrical circuit chapter because this instrument is valid and reliable with Cronbach's alpha reliability coefficient of 0.69. Furthermore, this physics assessment instrument in the form of a three-tier test can also be used by teachers to monitor the progress or effectiveness of classroom learning activities. The example of a test instrument developed by Peşman and Eryilmaz can be shown in Figure 6.



Directions:

Selection tool for text and ir

1. There are 12 questions. Please, try to answer each question carefully.
2. The batteries and the bulbs used in this test are identical.
3. The batteries are ideal, that is, the internal resistances are negligible.
4. Also, connection wires have got no resistance
5. If you do not find your answer among alternatives in any second tier, use the blank one for your response.

QUESTIONS

1.1. Will the bulb in Figure 1 light?
a) Yes, it will.
b) No, it will not.

1.2. Which one of the followings is the reason of your answer to the previous question?
a) The battery and the bulb are connected by the wire.
b) An extra wire must be connected from the negative terminal of the battery to the screw base of the bulb so that the positive and negative charges meet in the bulb.
c) An extra wire must be connected from the negative terminal of the battery to the screw base of the bulb so that the electric current passes through the bulb.
d) .....
.....

1.3. Are you sure about your answers given to the previous two questions?
a) Sure.
b) Not sure.

Figure 1

Figure 6.
Example of a Test Instrument Developed by Peşman and Eryilmaz

According to Yusrizal and Halim (2017) state that there are various forms of PysTT that are used to measure students' conceptual understanding and creative thinking skills in physics subjects. PysTT are developed in the form of one-tier, two-tier, and three-tier tests. The one-tier test developed has 8 items with 5 multiple-choice answers, while the two-tier test consists of 8 items with 5 answer choices plus a Contanity Response of Indexs (CRI) with a scale of 0.5, and a three-tier test consisting of 8 items, 5 answer choices, CRI index, and students are asked to write the reason for the answer. Based on the three forms of PysTT, misconceptions in students can be measured or detected clearly using a test instrument in the form of a three-tier test. This is because the three-tier test has the choice of the beliefs of the answers chosen by the students.

According to the expert's opinion and research results, it can be concluded that the assessment instrument in the form of a PysTT is a written test instrument in the form of multiple choices-tier developed to measure or assess all aspects of the ability of students, especially the ability to conceptual understanding and creative thinking skills on physic. Furthermore, PysTT are physics tests in the form of multiple choice questions that do not only require answers from students to a physics problem, but also require students to give reasons for their chosen answers. The number of tier-test that contain a choice of answers and reasons for a physics problem depends on the form of the PysTT, there are second physics-tier tests, third physics-tier tests, and four physics-tier tests. All forms of PysTT are effective for measuring students' conceptual understanding and creative thinking skills, except that in practice they need to be done wisely by considering several things, i.e the certainty of students' choices, time to investigate misconceptions, and ability of students to find reasons.

The Basic Theory of Conceptual Understanding

According to Kurniawati et al. (2017) state that learning concepts is the main outcome in the education process, the concept is a building block in thinking. The concept is the basis for higher mental processes to formulate principles and generalizations of a problem or event that occurs in life. In addition, the concept can also be interpreted as a set of meanings that contain order, patterns of relationships between objects and events. Conceptual understanding is the ability of students to understand and interpret objects, events, and phenomena as a whole using their own language construction (Moran and Keeley, 2015; Nugroho and Suryadarma, 2018). Students' conceptual understanding can be measured by indicators of students' abilities in translational understanding, interpretation, and extrapolation (Geary et al., 2017; Nugroho and Suryadarma, 2018). In addition to these indicators, conceptual understanding can also be explained in more detail, i.e:

1) Able to restate a concept, meaning students are able to restate the purpose of learning.
2) Able to classify objects according to certain characteristics according to the concept, meaning students are able to group objects according to their type and nature.
3) Able to distinguish between examples and not examples, meaning students are able to provide examples in everyday life in accordance with the concept of physics using their own language.
4) Able to present concepts in various forms of presentation, meaning that students are able to present or present learning material in various forms, whether in writing, graphics, or tables.
5) Able to develop necessary requirements and sufficient conditions, meaning that students are able to manage well in understanding and solving physics problems based on the boundaries of a learning material concept.
6) Able to use, utilize, and choose the right steps in solving problem solving problems, meaning that students are able to manage to solve a problem that occurs in life using creative, practical, and easy ways to do it based on the concept of learning material that they understand.
7) Able to classify problem solving concepts, meaning that students are able to manage, classify, and properly use the concepts of learning materials that are suitable to be used to solve problems that occur (Moran and Keeley, 2015).

According to Holme et al. (2015) state that conceptual understanding in contexts is the ability of students to understand the relationship of concepts to each other so that they can be applied to solve problems. Conceptual understanding in physics includes the ability to represent and translate physics problems in the form of macroscopic (observable) representations, microscopic (particles), graphics and so on simultaneously. Submission of several physics concepts that tend to be abstract

is very difficult to visualize in verbal form, so that requires the ability of the teacher to organize the content of lessons that can stimulate the process as preparation to build student knowledge, for example physics concepts that require deeper understanding because they tend to be abstract the concept of atomic energy (Faizah et al., 2013).

According to Eggen and Kauchak (2012), student knowledge and understanding of a subject matter concept can be measured in four ways, i.e students are able to define the concept of subject matter correctly; able to correctly identify the characteristics of a subject matter concept; able to connect concepts with other concepts appropriately, and be able to identify or provide examples of concepts that have never been encountered before. Therefore, students who have the right understanding of the concept mean that the students understand correctly about an abstract idea or concept that is being learned.

Therefore, it can be concluded that students' conceptual understanding is students' thinking in conceptual understanding especially physics so that they can restate the concept, classify objects according to certain properties, provide examples and not examples of the physics concepts they have studied, present concepts of physics in various representations, using certain procedures and applying their concepts to solving problems of daily life and in the process of learning physics.

The Basic Theory of Creative Thinking Skills

According to Birgili (2015) that the source of creativity is the tendency to actualize themselves, realize potential, drive to develop and mature, the tendency to express and activate themselves. Meanwhile, creative thinking is a thinking process that is oriented to a good and correct answer that needs to be trained to students because it can help students respond to a problem from various perspectives and are able to produce many ideas in solving the problem (Kusumaningrum and Djukri, 2016; Perry and Karpova, 2017). Furthermore, Nuswowati et al. (2017) suggests that creative thinking has four indicators, i.e fluence is the ability to produce many ideas, flexibility is the ability to produce varied ideas, originality, is the ability to produce original new ideas, and elaboration is the ability develop or add ideas to produce more detailed and innovative ideas.

Meanwhile, Perry and Karpova (2017) also stated that indicators or characteristics for people who have creative thinking skills, especially those related to physics, they can be seen or measured from the indications below, i.e

1) Fluency thinking skills, i.e skills to spark lots of ideas, answers/questions, problem solving, provide many suggestions for doing various things, and always think of more than one answer.

2) Flexible thinking skills (flexibility), i.e skills to generate ideas, answers/questions that vary; can see a problem from different perspectives; look for many different alternatives, and be able to change the way you think.
3) Original thinking skills (orginality), i.e the skill to be able to give birth to new expressions, think of common ways to express themselves, and be able to make unusual combinations of elements.
4) Elaboration skills, i.e skills to develop an idea, add or specify the details of an object, idea, or situation so that it becomes more interesting.
5) Assessing skills (evaluating), i.e the skills to determine the standard of self-assessment and determine whether a statement is true and able to make decisions on open situations not only spark ideas, but also implement them.
6) Redefinition skills, as the ability to review a problem through ways and perspectives that are different from what is already common.
7) Sensitivity skills in thinking (sensitivity), i.e the skill of a person in observing sensitively to capture and produce problems in response to a situation.

Creative thinking skills are a whole set of cognitive aspects skills that are used by each student according to the object, certain problems and conditions, or the type of effort towards certain events and problems based on the student's capacity. In general, creative thinking skills are closely related to critical thinking skills, and problem solving skills, because all three are part of high-order thinking skills (HOTS) that must be controlled by each student (Muskitta and Djukri, 2016). Actually, there are three dimensions of creative thinking skills, i.e

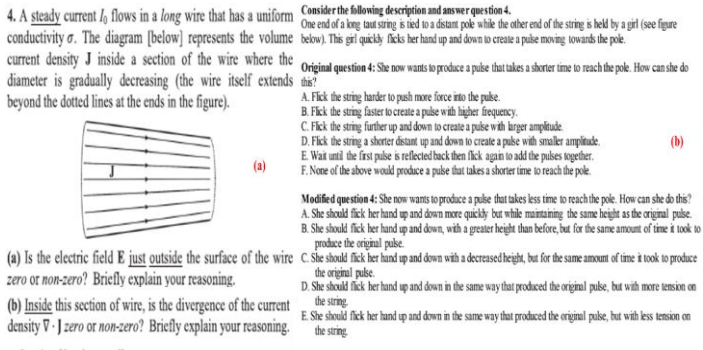
1) Synthesis, covering various activities such as getting benefits from analogical thinking, deducing from small or simple parts, presenting new and informal suggestions in solving a problem.
2) Articulation, involves forming old and new knowledge or expanding and integrating current knowledge with new thoughts or challenges, and creating unusual relationships to produce informal solutions.
3) Imagination, includes activities that create a connection between valid and reliable thinking and presents a flexible way of thinking with the help of imagination to produce varied insights during the process of making ideas.

Furthermore, creative thinking skills are also skills possessed by each individual to look for new ways, strategies, ideas or ideas how to obtain solutions to a problem at hand (Santofani and Rosana, 2016). In addition, creative thinking is the ability to think that starts from the sensitivity of the situation at hand, that the situation is seen or identified as a problem that wants to or must be resolved. This way of thinking is needed in studying physics, because physics has a strong and clear structure and linkages between concepts so that students are accustomed to using the above skills in developing physics creative thinking skills when students are in problem solving (Wibowo and Suhandi, 2013).

Therefore, based on the expert's view of creative thinking skills, it can be concluded that creative thinking skills are the skills of using creative ideas and techniques that are broadly unlimited; creating useful new ideas; describe, reconcile, analyze, and evaluate existing ideas in order to develop and maximize creative efforts. Furthermore, in the process of creative thinking requires an attitude of openness, courage to take risks, tolerance to differences, and self-discipline.

The Concept of Conceptual Understanding Physics Test

The following are several sources of research that have been carried out by several related experts in using various types of tests to measure students' conceptual understanding about physics matter. According to Baily et al. (2017) stated that assessment instruments were developed to measure students' conceptual understanding in physics material especially the electrodynamic chapter in the form of essay questions or free responses consisting of 6 questions with each question having 15 sub-questions, which focusing on students' conceptual understanding of electrodynamic material. Research conducted by Barniol and Zavala (2016) it is the development of a one-tier test instrument, in this case only developing tests that demand student answers freely in accordance with the physics concept of the electrodynamics chapter. In addition, students' conceptual understanding of electrodynamic material which is assessed using one-tier test assessment instruments is quite high as evidenced by students being able to interpret questions as intended and the total value of test results correlating well with other variables, such as final examinations and subject scores physics. The example of a test instrument developed by Baily, Barniol, and Zavala can be shown in Figure 7.



4. A steady current I_0 flows in a *long* wire that has a uniform conductivity σ. The diagram [below] represents the volume current density **J** inside a section of the wire where the diameter is gradually decreasing (the wire itself extends beyond the dotted lines at the ends in the figure).

(a)

(a) Is the electric field **E** just outside the surface of the wire *zero* or *non-zero*? Briefly explain your reasoning.

(b) Inside this section of wire, is the divergence of the current density $\nabla \cdot \mathbf{J}$ *zero* or *non-zero*? Briefly explain your reasoning.

Consider the following description and answer question 4.
One end of a long taut string is tied to a distant pole while the other end of the string is held by a girl (see figure below). This girl quickly flicks her hand up and down to create a pulse moving towards the pole.

Original question 4: She now wants to produce a pulse that takes a shorter time to reach the pole. How can she do this?
A. Flick the string harder to push more force into the pulse.
B. Flick the string faster to create a pulse with higher frequency.
C. Flick the string further up and down to create a pulse with larger amplitude.
D. Flick the string a shorter distant up and down to create a pulse with smaller amplitude.
E. Wait until the first pulse is reflected back then flick again to add the pulses together.
F. None of the above would produce a pulse that takes a shorter time to reach the pole.

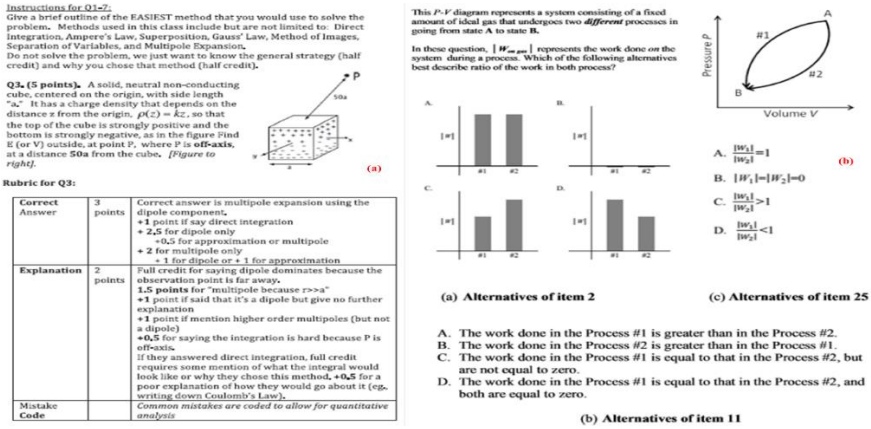
(b)

Modified question 4: She now wants to produce a pulse that takes less time to reach the pole. How can she do this?
A. She should flick her hand up and down more quickly but while maintaining the same height as the original pulse.
B. She should flick her hand up and down, with a greater height than before, but for the same amount of time it took to produce the original pulse.
C. She should flick her hand up and down with a decreased height, but for the same amount of time it took to produce the original pulse.
D. She should flick her hand up and down in the same way that produced the original pulse, but with more tension on the string.
E. She should flick her hand up and down in the same way that produced the original pulse, but with less tension on the string.

Figure 7.
Example of a Test Instrument Developed by, a) Baily and b) Barniol and Zavala

According to Chasteen et al. (2012) state that students' conceptual understanding of physics material in electrostatics can be measured effectively using a two-tier multiple choice test instrument. This two-tier test asks students to choose the problem solving method from the questions provided (one tier test) and explain the problem solving steps that have been chosen (second tier test). Students' conceptual

understanding of physics material besides being able to be measured using tier multiple choice test instruments can also be measured using tier tests that require various kinds of representations of answers to physics problems. Sriyansyah and Suhandi (2016) stated that test instruments used to measure students' understanding of physics material can also be carried out effectively using test instruments that demand various forms of student answers. This assessment instrument is in the form of a multiple choice of 30 items with three different representations/answer groups, i.e verbal, mathematical, and diagram. The example of a test instrument developed by Chasteen, Sriyansyah, and Suhandi can be shown in Figure 8.



Instructions for Q1-7:
Give a brief outline of the EASIEST method that you would use to solve the problem. Methods used in this class include but are not limited to: Direct Integration, Ampere's Law, Superposition, Gauss' Law, Method of Images, Separation of Variables, and Multipole Expansion.
Do not solve the problem, we just want to know the general strategy (half credit) and why you chose that method (half credit).

Q3. (5 points). A solid, neutral non-conducting cube, centered on the origin, with side length "a." It has a charge density that depends on the distance z from the origin, $\rho(z) = kz$, so that the top of the cube is strongly positive and the bottom is strongly negative, as in the figure Find E (or V) outside, at point P, where P is **off-axis**, at a distance **50a** from the cube. *[Figure to right].*

(a)

Rubric for Q3:

Correct Answer	3 points	Correct answer is multipole expansion using the dipole component. **+1** point if say direct integration **+ 2.5** for dipole only +0.5 for approximation or multipole **+ 2** for multipole only + 1 for dipole or + 1 for approximation
Explanation	2 points	Full credit for saying dipole dominates because the observation point is far away. **1.5 points** for "multipole because r>>a" **+1** point if said that it's a dipole but give no further explanation **+1** point if mention higher order multipoles (but not a dipole) **+0.5** for saying the integration is hard because P is off-axis. If they answered direct integration, full credit requires some mention of what the integral would look like or why they chose this method. **+0.5** for a poor explanation of how they would go about it (eg. writing down Coulomb's Law).
Mistake Code		*Common mistakes are coded to allow for quantitative analysis*

This *P-V* diagram represents a system consisting of a fixed amount of ideal gas that undergoes two ***different*** processes in going from state A to state B.

In these question, $|W_{on\ gas}|$ represents the work done *on* the system during a process. Which of the following alternatives best describe ratio of the work in both process?

A. $\frac{|W_1|}{|W_2|}=1$
B. $|W_1|=|W_2|=0$
C. $\frac{|W_1|}{|W_2|}>1$
D. $\frac{|W_1|}{|W_2|}<1$

(b)

(a) **Alternatives of item 2** (c) **Alternatives of item 25**

A. The work done in the Process #1 is greater than in the Process #2.
B. The work done in the Process #2 is greater than in the Process #1.
C. The work done in the Process #1 is equal to that in the Process #2, but are not equal to zero.
D. The work done in the Process #1 is equal to that in the Process #2, and both are equal to zero.

(b) **Alternatives of item 11**

Figure 8.
Example of a Test Instrument Developed by, a) Chasteen and b) Sriyansyah, and Suhandi

According to Eshach (2014) also states that students' conceptual understanding can also be measured effectively using a three-tier multiple choice test instrument. The test instrument developed is one of the student-centered test instruments, students are asked to provide their own answers or choose answers in the form of the reason for a question. This test instrument developed by Eshach has a multiple choice assessment format that requires students to choose one answer in the form of an excuse and states whether the chosen reason is true or false. After that, students were also asked to determine their tier of confidence in the answers chosen on a scale of 1-5. The example of a test instrument developed by Eshach can be shown in Figure 9.

1. *Can there be a sound that we do not hear?*
   a. Yes. Our ears admit only sound particles of certain sizes. Animal ears admit different sizes of sound particles, so they can hear sounds we don't hear, and vice versa.
      True/False. Certainty level in answer: 1, 2, 3, 4, 5
   b. Yes. We can hear because our eardrum can detect changes in the movement of the air surrounding it. Our eardrum works in a certain range of air pressure.
      True/False. Certainty level in answer: 1, 2, 3, 4, 5
   c. None of the above choices fits my basic viewpoint. My basic viewpoint is (please explain your viewpoint in the space provided below):
2. *When we strum a guitar string, we hear a sound because:*
   a. The vibrating string releases sound particles and pushes them outward so they reach our ears.
      True/False. Certainty level in answer: 1, 2, 3, 4, 5

Figure 9.
Example of a Test Instrument Developed by Eshach

Therefore, based on the expert's view of the research they have done in measuring students' conceptual understanding of physics material, it can be concluded that students' conceptual understanding of physics learning material can be measured using a variety of test instruments, either multiple choice test or open-ended test. However, in general of students' conceptual understanding of physics is more effective when measured using a tier test instrument, both one-tier, two-tier, three-tier, four-tier multiple-choice tests, or tests that require a variety of student answers. This is because a physics tier-test instrument is a student-centered test instrument, students are asked to give their own answers or choose answers in the form of the reason for a question.

The Concept of Creative Thinking Skills Physics Test

The following are several sources of research that have been carried out by several related experts in using various types of tests to measure students' creative thinking skills about physics matter. According to Taylor and Getzels (1975) and Piaw (2010) states that the general criteria for choosing specific creative thinking skills tests in physics subjects are, must have relevance to good physics theory, must have relevance to creative thinking behavior in the real world, only aspects of thinking skills are different, must be attractive to respondents, especially students, must be built so that someone can respond. In terms of any of his experiences, this creative thinking skills test instrument must produce data that can be assessed reliably for aspects of thought, testing material, deadlines, and assessment procedures must be clearly stated and relevant.

According to Noviani et al. (2017) state that students' creative thinking skills can be measured using open-ended tests that ask students to solve physics problems. In addition, there are variations in the tests used to measure students' creative thinking

skills in physics subjects, i.e using parallel tests that almost resemble physics-tier tests (PysTT). Almeida et al. (2017) state that tests to measure students' creative thinking skills in physics subjects have two parallel forms, A and B, including the following subjects: (a) Asking questions and making guesses (subtests 1, 2, and 3), where students write questions and make guesses about the possible causes and consequences of the situation based on the images presented; (b) Product improvement (subtest 4); (c) Unusual use (subtest 5), where the list of students is interesting and it is not unusual to use a cardboard box; and (d) Consider (subtest 6), where students are asked to record all the consequences if a situation is not possible.

Meanwhile, the other parallel tests consist of two parallel forms with three subtests, i.e (a) compiling images of physics problems; (b) completing a images of physics problems; and (c) arrange images about different physics problems from parallel lines. Both forms of parallel tests aim to assess the four main cognitive processes of creativity: (a) fluency or the number of relevant responses; (b) flexibility as referred to in various categories; (c) authenticity requires relevant new considerations; and (d) the elaboration referred to in the number of details used to provide a response (Almeida et al., 2017). Furthermore, Karpova et al. (2015) stated that the assessment of creative thinking skills can be done effectively using a two-tier test instrument in the form of an open-ended test that requires students to answer questions in the figural and verbal format. The figural format was chosen to assess the quantity and quality of creative ideas triggered by each student in solving every physics problem.

Thus, based on several opinions expressed by experts in various studies, it can be concluded that the assessment to measure students' creative thinking skills in physics can be done using a variety of test instruments which are essentially able to provide opportunities for students to express their thoughts or ideas creative in completing various topics on physics problems. Moreover, the test instrument that is generally used is a physics tier test instrument that each question requires each student to provide the answer and the creative reasons for the answer or it can also require students to provide answers in various representations in the form of diagrams or writing that can accommodate each aspects of students' creative thinking skills, namely fluency, flexibility, originality, and elaboration.

Qualitative Synthesis of Physics-Tier Tests to Measure Students' Conceptual Understanding and Creative Thinking Skills

Based on some experts 'opinions on each concept in a physics-tier test instrument to measure students' conceptual understanding and creative thinking skills, the overall synthesis of this research can be shown in Table 2.

Table 2.
Synthesis of Physics-Tier Tests to Measure Students' Conceptual Understanding and Creative Thinking Skills

Aspect	Expert	Expert' Opinion or Research
Physics-Tier Tests (PysTT)	Winarti, Cari, Suparmi, Sunarno, and Istiyono	The two-tier test is a test instrument used to measure students' conceptual understanding in the form of reasoned multiple choice tests.
	Hermita, Suhandi, Syaodih, Samsudin, Isjoni, Johan, Rosa, Setyaningsih, and Sapriadi	The tests instruments used to measure conceptual understanding and creative thinking skills can use physics four-tier tests.
	Pes¸ Ma and Eryilmaz	Other test instruments that can be used to measure students' conceptual understanding and creative thinking skills are three-tier multiple choice tests.
	Yusrizal and Halim	There are various forms of physics-tier tests that are used to measure students' conceptual understanding and creative thinking skills in physics subjects.
Conceptual Understanding	Kurniawati, Hartanto, and Zamzaili	The learning concepts is the main outcome in the education process, the concept is a building block in thinking.
	Moran and Keeley & Nugroho and Suryadarma	Conceptual understanding is the ability of students to understand and interpret objects, events, and phenomena as a whole using their own language construction.
	Geary, vanMarle, Chu, Rouder, Hoard, and Nugent	Students' conceptual understanding can be measured by indicators of students' abilities in translational understanding, interpretation, and extrapolation.
	Holme, Luxford, and Brandriet	Conceptual understanding is the ability of students to understand the relationship of concepts to each other so that they can be applied to solve problems.
Creative Thinking Skills	Birgili	Source of creativity is the tendency to actualize themselves, realize potential, and drive to develop and mature.
	Perry and Karpova & Kusumaningrum and Djukri	Creative thinking is a thinking process that is oriented to a good and correct answer that needs to be trained to students because it can help students respond to a problem from various perspectives.
	Nuswowati, Susilaningsih, Ramlawati and Kadarwati	Creative thinking has four indicators, i.e fluence, flexibility, originality, and elaboration.
	Muskitta and Djukri	Creative thinking skills are part of high-order thinking skills (HOTS).
	Santofani and Rosana	Creative thinking skills are skills possessed by each individual to look for new ways, strategies, ideas how to obtain solutions to a problem at hand.
Conceptual Understanding Physics Test	Baily, Ryan, Astolfi, and Pollock	The assessment instruments in the form of essay questions or free responses can be used to measure students' conceptual understanding.

Aspect	Expert	Expert' Opinion or Research
	Barniol and Zavala	The one-tier test instrument, in this case only developing tests that demand student answers freely in accordance with the physics concept.
	Chasteen, Pepper, Caballero, Pollock, and Perkins	The students' conceptual understanding of physics material can be measured effectively using a two-tier multiple choice test instrument.
	Sriyansyah and Suhandi	Test instruments used to measure students' understanding of physics material can also be using test instruments that demand various forms of student answers.
	Eshach	The students' conceptual understanding can be measured using a three-tier multiple choice test instrument.
Creative Thinking Skills Physics Test	Noviani, Hartono, and Rusilowati	The students' creative thinking skills can be measured using open-ended tests.
	Almeida, Prietob, Ferrando, Oliveira, and Ferrandiz	Tests to measure students' creative thinking skills in physics subjects have two parallel forms, A and B.
	Karpova, Marcketti, and Barker	The assessment of creative thinking skills can be done using a two-tier test instrument in the form of an open-ended test that requires students to answer questions in the figural and verbal format.
Synthesis		
An assessment instrument in the form of a physics-tier test (PysTT) is a written test instrument in the form of multiple-choice tier developed to measure conceptual understanding which emphasizes re-expressing physics concepts in various representations and creative thinking skills in physics which includes fluency, flexibility, originality, and elaboration.		

Validity and Reliability Results for Reasonable Multiple Choice Test Questions or Physics Two-Tier Tests

After we find out together about what data is sought in this study using instruments that have been developed by researchers, we first need to know about the feasibility of research instruments that researchers have developed. The feasibility data of this research instrument is in the form of data about the validity and reliability of the instrument that has been provided by experts and validator practitioners, and has been analyzed using equations 1 and 2. The first analysis of the research instrument is to analyze the feasibility (validity and reliability) of reasoned multiple choice test questions or physics two-tier test. Meanwhile, the results of the validity analysis of reasoned multiple choice tests or physics two-tier test can be shown in Table 3 below.

Table 3.
Validation Results of Reasoned Multiple Choice Tests or Physics Two-Tier Test

Assessment Item	Validity (V)	Category
Creative Thinking Skills		
Aspects of Learning Guides	0.84	Good
Quality Aspects of Matter in Test	0.70	Good
Aspects of Pictures and Language	0.85	Very Good
Aspects of Conformity of Creative Thinking Skills Test	0.77	Good
Validity of Creative Thinking Skills Test	**0.79**	**Good**
Conceptual Understanding		
Aspects of Learning Guides	0.87	Good
Quality Aspects of Matter in Test	0.79	Good
Aspects of Pictures and Language	0.89	Very Good
Aspects of Conformity of Conceptual Understanding Test	0.76	Good
Validity of conceptual understandingTest	**0.83**	**Very Good**
Validity of Reasoned Multiple Choice Tests or Physics Two-Tier Test	**0.81**	**Very Good**

Based on the results of the validation of the instruments used in this study which can be shown in Table 3, the validity results of reasoned multiple choice tests or physics two-tier tests were 0.81 with very good categories. With the breakdown, the result of the validity of the physics two-tier tests of creative thinking skills is 0.79 with a good category and the result of the validity of the physics two-tier tests of the ability to conceptual understanding is 0.83 which is included in the very good category. Therefore, it can be stated that the reasoned multiple choice tests or physics two-tier tests are valid and suitable for measuring creative thinking skills and conceptual understanding of 10th grade students at SMA N 9 Yogyakarta.

After the reasoned multiple choice tests or physics two-tier tests used in this study were analyzed for validity using the Aiken V equation, the next step is to analyze the other parts of the feasibility, namely analyzing the reliability of reasoned multiple choice or physics two-tier tests. In other words, the results of the reliability are also used as part of the feasibility of a multiple choice test or physics two-tier test that has been developed by the researchers. The results of the reliability of reasoned multiple choice tests or physics two-tier tests can be presented in Table 4 below.

Table 4.

Reliability of Reasoned Multiple Choice Tests or Physics Two-Tier Tests

Reliability of Creative Thinking Skills Test		
Item number of Creative Thinking Skills Test	**Average of PA per Item (%)**	**Category**
1.	91.33	Reliable
2.	96.32	Reliable
3.	98.65	Reliable
4.	97.75	Reliable
5.	96.21	Reliable
6.	98.26	Reliable
7.	98.16	Reliable
8.	97.37	Reliable
9.	96.46	Reliable
10.	97.55	Reliable
11.	97.66	Reliable
12.	97.74	Reliable
13.	98.45	Reliable
14.	98.55	Reliable
15.	98.75	Reliable
16.	99.27	Reliable
17.	99.16	Reliable
18.	99.08	Reliable
19.	94.87	Reliable
20.	95.97	Reliable
21.	96.66	Reliable
22.	97.05	Reliable
23.	96.64	Reliable
24.	98.77	Reliable
25.	98.89	Reliable
Quantity	**98.72**	**Reliable**
Reliability of Conceptual Understanding Test		
Item number of Conceptual Understanding Test	**Average of PA per Item (%)**	**Category**
1.	98.43	Reliable
2.	97.37	Reliable
3.	98.65	Reliable
4.	97.75	Reliable
5.	98.27	Reliable
6.	99.24	Reliable
7.	99.46	Reliable
8.	96.97	Reliable
9.	98.76	Reliable
10.	94.55	Reliable
11.	98.66	Reliable
12.	97.84	Reliable
13.	96.65	Reliable
14.	98.45	Reliable
15.	99.75	Reliable
16.	96.77	Reliable
17.	98.66	Reliable
18.	98.18	Reliable
19.	99.37	Reliable
20.	97.57	Reliable
21.	98.56	Reliable
22.	97.95	Reliable
23.	96.74	Reliable
24.	99.57	Reliable
25.	98.87	Reliable
Quantity	**98.92**	**Reliable**
Reliability of Reasoned Multiple Choice Tests or Physics Two-Tier Tests	**98.82**	**Reliable**

Based on Table 4, we can observe that in general the reasoned multiple choice tests or physics two-tier tests that have been developed by researchers are reliable with a reliability percentage of 98.82. Meanwhile, in detail, we can see empirical data on the reliability level of reasoned multiple choice tests or physics two-tier tests of creative thinking skills obtained percentage reliability of 98.72% and 98.92% for the ability to conceptual understanding. The detailed data are all categorized as reliable because they meet the reliability requirements of the percentage agreement value (PA), where the research instrument in this case reasoned multiple choice tests or physics two-tier tests can be said to be reliable, if the percentage agreement value (PA) obtained for each research instrument is more than 75% (Borich, 1994).

Based on Table 4, we can observe that the reliability results of reasoned multiple choice tests or physics two-tier tests obtain different reliability results. In this case the reliability level of reasoned multiple choice tests or physics two-tier tests for the ability to conceptual understanding obtain higher reliability results than the reliability of creative thinking skills. However, the reliability of reasoned multiple choice tests or physics two-tier tests are at 90% intervals and all fall into the reliable category. This can happen, one of which is influenced by the assessor who gives the lowest and highest score of many evaluators. In addition, it can also be caused by the achievement of creative thinking skills test results are lower than has is the ability to conceptual understanding, so the percentage of reliability is also higher for the ability to conceptual understanding. This often happens because each assessor gives an assessment of each research instrument that has a different view, there are times when the first assessor gives the highest score on a particular item number, but other assessors give the lowest score on the item number (Yuliani & Saragih, 2015). Therefore, the percentage of reliability gain of each reasoned multiple choice tests or physics two-tier tests for creative thinking skills and conceptual understanding are also different, but still in the same interval.

The Results of Creative Thinking Skills and Conceptual Understanding

After obtaining the validity and reliability results on the questions of creative thinking skills and conceptual understanding, then discussing the results of creative thinking skills and conceptual understanding of 10th grade students in SMA N 9 Yogyakarta that can be measured using the assessment instruments. The results of creative thinking and conceptual understanding of 10th grade students at SMA N 9 Yogyakarta can be observed in the following Figure 10.

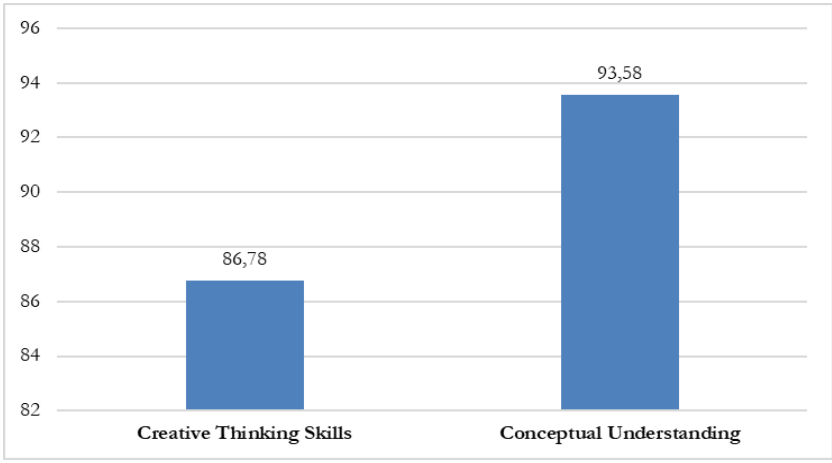


Figure 10.
Results of the Value of Creative Thinking Skills and Students' Conceptual Understanding

Based on Figure 10, which shows that the results of creative thinking skills of 10th grade students at SMA N 9 Yogyakarta is lower than the ability to understand their physical concepts. Difference in the acquisition of the average value of creative thinking skills and understanding of the concept of 10th grade students at SMA N 9 Yogyakarta by 6.8 or 0.068%. However, the results of creative thinking skills and understanding of the concepts of 10th grade students at SMA N 9 Yogyakarta are included in the good category. This is because the value of the two aspects is more than the value of the minimum completeness criteria (KKM) set at SMA N 9 Yogyakarta of 75. In addition, these results are also caused by the cognitive development characteristics of middle school students who begin at the formal operational stage who begin to have creative ideas (Asyari, Al Muhdhar, Susilo, & Ibrohim, 2016). Therefore, it is easier for them to understand the concepts of physics conveyed by the teacher than it is for creative thinking skills. In addition, these results also show some errors that occur during physics learning activities, one of which teachers still rarely provide variations of physics practice questions in addition to calculations and the tendency of learning physics is only centered on the teacher (Fuad, Zubaidah, Mahanal, & Suarsini, 2017).

Conclusion

Based on the results of this study, it can be concluded that an assessment instrument in the form of a physics-tier test (PysTT) is a written test instrument in the form of multiple-choice tier developed to measure or assess all aspects of a student's ability. The ability of these students specifically conceptual understanding which emphasizes re-expressing physics concepts in various representations, detecting physics misconceptions that occur in some students, solving problems that occur in life based on the correct concepts of physics; and creative thinking skills in physics

which includes fluency, flexibility, originality, and elaboration in solving any physics problems that occur in life using solutions that are creative, innovative, and adaptive. Furthermore, the physics-tier test (PysTT) is a physics test in the form of multiple choice questions that not only requires answers from students for physics problems, but also requires students to provide reasons for their chosen answers. This test prioritizes the reasons given by students, so students can provide solutions according to their creativity and innovation based on the physics concepts they know. Therefore, this test is precisely used to measure whether students have understood the concepts of physics correctly and whether students experience misconceptions or not. All forms of physics-tier tests (PysTT) whether in the form of one-tier tests, two-tier, three-tier, or even four-tier test can measure students' conceptual understanding and creative thinking skills, but in practice they must be done wisely and adapted to students' ability to physics and test time. In addition, the results of the validity and reliability of reasoned multiple choice test or physics two-tier tests to measure students' conceptual understanding and creative thinking skills were 0.81 and 98.82%, which were included in the very good and reliable category. The ability to conceptual understanding and creative thinking skills of students who are measured using the physics two-tier test is very good because it is more than the value of the specified minimum completeness criteria. Even so, the results of students' creative thinking skills are lower than the ability to understand their concepts. This is in line with the research of Walid, Sajidan, Ramli, and Kusumah (2019) which states that creative thinking skills are part of higher order thinking skills (HOTS) which tend to be weaker among students than other abilities.

Acknowledgements

We would like to thank the lecturers of the physics education program of the Yogyakarta State University postgraduate program who have provided moral and material guidance and assistance in the preparation of this article.

Biodata of the Author

Himawan Putranta is a M.Pd. Candidate, Department of Physics Education, Graduate School of Yogyakarta State University, Colombo Street No.1 Yogyakarta, Yogyakarta 55281, Indonesia.
E-mail: himawanputranta.2017@student.uny.ac.id
Phone: +6285742977623
ORCID ID: https://orcid.org/0000-0003-2646-8237

Supahar is a M.Si., Dr., Department of Physics Education, Graduate School of Yogyakarta State University, Colombo Street No.1 Yogyakarta, Indonesia.
E-mail: supahar@uny.ac.id
Phone: +628164267452
ORCID ID: https://orcid.org/0000-0002-2486-5549

References

Aizikovitsh-Udi, E., & Cheng, D. (2015). Developing critical thinking skills from dispositions to abilities: Mathematics education from early childhood to high school. *Creative Education, 6*(1), 455-462.

Almeida, L. S., Prietob, L. P., Ferrando, M., Oliveira, E., & Ferrandiz, C. (2017). Torrance Test of Creative Thinking: The question of its construct validity. *Journal of Innovative Science Education, 3*(1), 53-58.

Apino, E., & Retnawati, H. (2017). Developing instructional design to improve mathematical higher order thinking skills of students. *Journal of Physics: Conference Series, 812*, 1-8.

Asyari, M., Al Muhdhar, M. H. I., Susilo, H., & Ibrohim. (2016). Improving critical thinking skills through the integration of problem based learning and group investigation. *International Journal for Lesson and Learning Studies*, *5*(1), 36-44.

Azwar, S. (2015). *Reliability and validity of edition 4*. Yogyakarta: Pustaka Pelajar.

Baily, C., Ryan, Q. X., Astolfi, C., & Pollock, S. J. (2017). Conceptual assessment tool for advanced undergraduate electrodynamics. *Physical Review Physics Education Research*, *13*(2), 020113.

Barniol, P., & Zavala, G. (2016). Mechanical waves conceptual survey: Its modification and conversion to a standard multiple-choice test. *Physical Review Physics Education Research*, *12*(1), 010107.

Barrow, R. (2015). *Understanding skills: Thinking, feeling, and caring 1st edition*. London: Routledge.

Bates, S. P., & Galloway, R. K., Riise, J., & Homer, D. (2014). Assessing the quality of a student-generated question repository. *Physical Review Physics Education Research, 10*(2), 020105.

Birgili, B. (2015). Creative and critical thinking skills in problem-based learning environments. *Journal of Gifted Education and Creativity*, 2(2), 71-80.

Borich, Gary D. (1994). *Observation skill for effective teaching.* New York: Macmillan Publishing Company.

Caballero, M. D., Doughty, L., Turnbull, A. M., Pepper, R. E., & Pollock, S. J. (2017). Assessing learning outcomes in middle-division classical mechanics: The Colorado classical mechanics and math methods instrument. *Physical Review Physics Education Research*, 13(1), 010118.

Chasteen, S. V., Pepper, R. E., Caballero, M. D., Pollock, S. J., & Perkins, K. K. (2012). Colorado Upper-Division Electrostatics diagnostic: A conceptual assessment for the junior level. *Physical Review Physics Education Research. 8*(2), 020108.

Collins, R. (2014). Skills for the 21st Century: Teaching higher-order thinking. *Curriculum & Leadership Journal, 12*(14), 1-7.

Ding, L. (2014). Seeking missing pieces in science concept assessments: Reevaluating the brief electricity and magnetism assessment through rasch analysis. *Physical Review Physics Education Research, 10*(1), 010105.

Docktor, J. L., Dornfeld, J., Frodermann, E., Heller, K., & Hsu, L. (2016). Assessing student written problem solutions: A problem-solving rubric with application to introductory physics. *Physical Review Physics Education Research, 12*(1), 010130.

Efendi, R. (2010). Kemampuan fisika siswa indonesia dalam TIMSS. *Prosiding Seminar Nasional Fisika 2010, ISBN: 978-979-98010-6-7*.

Eshach, H. (2014). Development of a student centered instrument to assess middle school students' conceptual understanding of sound. *Physical Review Physics Education Research, 10*(1), 010102.

Faizah, Miswadi, S. S., & Haryani, S. (2013). Development of problem-based learning devices to improve soft skills and understanding concept. *Jurnal Pendidikan IPA Indonesia, 2*(2), 120-128.

Fuad, N. M., Zubaidah, S., Mahanal, S., & Suarsini, E. (2017). Improving junior high schools' critical thinking skills based on test three different models of learning. *International Journal of Instruction*, *10*(1), 101-116.

Furberg, A. (2016). Teacher support in computer-supported lab work: Bridging the gap between lab experiments and students' conceptual understanding. *International Journal of Computer-Supported Collaborative Learning, 2*(1), 1-25.

Geary, D. C., vanMarle, K., Chu, F. W., Rouder, J., Hoard, M. K., & Nugent, L. (2017). Early conceptual understanding of cardinality predicts superior school-entry number-system knowledge. *Psychological Science, 1*(2), 1-15.

Getzels, J. W., & Taylor, I. A. (Eds.). (1975). *Perspectives in creativity*. AldineTransaction.

Grbich, C. (2013). *Qualitative data analysis: An introduction*. London, England: Sage.

Gruenewald, D. A., & Smith, G. A. (2014). *Place-based education in the global age: Local diversity*. New York: Psychology Press.

Hermita, N., Suhandi, A., Syaodih, E., Samsudin, A., Isjoni, Johan, H., Rosa, F., Setyaningsih, R., Sapriadil, & Safitri, D. (2017). Constructing and implementing a four tier test about static electricity to diagnose pre-service elementary school teacher' misconceptions. *Journal of Physics: Conference Series*, 895.

Holme, T. A., Luxford, C. J., & Brandriet, A. (2015). Defining conceptual understanding in general chemistry. *Journal of Chemical Education*, *92*(9), 1477-1483.

Istiyono, E., Mardapi, D., & Suparno. (2014). Development of high physical thinking skills (PysTHOTS) development students in high school. *Jurnal Penelitian dan Evaluasi Pendidikan, 18*(1), 1-12.

Karpova, E., Marcketti, S. B., & Barker, J. (2011). The efficacy of teaching creativity: Assessment of student creative thinking before and after exercises. *Clothing and Textiles Research Journal, 29*(1), 52-66.

Kauchak, D., & Eggen, P. (2012). *Learning strategies and models: Teaching content and thinking skills*. Jakarta: Indeks.

Kiryak, Z., & Çalik, M. (2017). Improving grade 7 students' conceptual understanding of water pollution via common knowledge construction model. *International Journal of Science and Mathematics Education, 1*(1), 1-22.

Kurniawati, E., Hartanto, & Zamzaili. (2017). The effect of assurance, relevance, interest, assessment, satisfaction (ARIAS) integrative and early ability model in improving the understanding of mathematical concept and troubleshooting the first middle school students in Kepahiang. *Jurnal Pendidikan Matematika Raflesia, 2*(2), 174-187.

Kusumaningrum, S., & Djukri, D. (2016). Development of project based learning (PjBL) learning tools to improve skills in the process of science and creativity. *Jurnal Inovasi Pendidikan IPA, 2*(2), 241-251.

Loewenthal, K., & Lewis, C. A. (2018). *An introduction to psychological tests and scales*. London: Psychology Press.

Luangrath, P., Pettersson, S., & Benckert, S. (2011). On the use of two versions of the Force Concept Inventory to test conceptual understanding of mechanics in Lao PDR. *Eurasia Journal of Mathematics, Science and Technology Education*, 7(2), 103-114.

Mądrala, M., Wąsik, M., & Małoszewski, P. (2017). Interpretation of environmental tracer data for conceptual understanding of groundwater flow: An application for fractured aquifer systems in the Kłodzko Basin, Sudetes, Poland. *Isotopes in Environmental and Health Studies, 2*(1), 1-19.

Marshall, C., & Rossman, G. B. (2014). *Designing qualitative research (6th ed)*. London, England: Sage.

Moran, K. R., & Keeley, P. (2015). *Teaching for conceptual understanding in science*. Arlington: NSTA Press, National Science Teachers Association.

Muskitta, M., & Djukri. (2016). Effect of PBT model on critical thinking ability and creative thinking ability of high school students. *Jurnal Inovasi Pendidikan IPA, 2*(1), 58-65.

Neuman, W. L. (2011). *Social research methods: Qualitative and quantitative approaches*. Boston: Allyn & Bacon.

Newcombe, N. S., & Shipley, T. F. (2014). *Studying visual and spatial reasoning for design creativity: Thinking about spatial thinking*. Dordrecht: Springer.

Noviani, Y., Hartono, & Rusilowati, A. (2017). Analysis of students' mindset in solving science questions in terms of critical and creative thinking ability and science literacy. *Journal of Innovative Science Education, 6*(2), 147-154.

Nugroho, R. A., & Suryadarma, I. G. P. (2018). The influence of servant leadership learning with a concept mapping of students' understanding of concepts and critical thinking. *Jurnal Inovasi Pendidikan IPA, 4*(1), 114-127.

Nuswowati, M., Susilaningsih, E., Ramlawati, & Kadarwati, S. (2017). Implementation of problem-based learning with green chemistry vision to improve creative thinking skill and students' creative actions. *Jurnal Pendidikan IPA Indonesia, 6*(2), 221-228.

Perry, A., & Karpova, E. (2017). Efficacy of teaching creative thinking skills: A comparison of multiple creativity assessments. *Thinking Skills and Creativity, 1*(2), 1-25.

Peşman, H., & Eryilmaz, A. (2010). Development of a three tier test to assess misconceptions about simple electric circuits. *The Journal of Educational Research, 103*, 208–222.

Piaw, C. Y. (2010). Building a test to assess creative and critical thinking simultaneously. *Procedia Social and Behavioral Sciences, 2*(2), 551-559.

Porter, M. E., & Kramer, M. R. (2018). *Managing sustainable business: Creating shared value*. Dordrecht: Springer.

Rahmatan, H., Liliasari, & Redjeki, S. (2012). Development of computer-based biochemical learning models to collect creative thinking skills of biological teacher students. *Jurnal Pendidikan IPA Indonesia, 1*(2), 178-182.

Sadiqin, I. K., Santoso, U. T., & Sholahuddin, A. (2017). Understanding the natural science concept of middle school students through learning problem solving on topics of changing objects around us. *Jurnal Inovasi Pendidikan IPA, 3*(1), 52-62.

Santofani, A., & Rosana, D. (2016). Development of creativity tests on physics learning with inquiry approaches to the kinetic theory of gas material. *Jurnal Inovasi Pendidikan IPA, 2*(2), 134-144.

Sriyansyah, S. P., & Suhandi, A. (2016). Development of a representational conceptual evaluation in the first law of thermodynamics. *Journal of Physics: Conference Series 739*, 012125.

Sudjana, N. (1990). *Assessment of teaching and learning outcomes*. Bandung: PT Remaja Rosdakarya.

TIMSS & PIRLS International Study Center. (2012). TIMSS 2011 international results in science. Boston: The TIMSS & PIRLS International Study Center, Boston College.

Van den Berg, G. (2008). The use of assess-ment in the development of higher-order thinking skills. *Africa Education Review, 1*(2), 279-294.

Walid, A., Sajidan, S., Ramli, M., & Kusumah, R. G. T. (2019). Construction of the assessment concept to measure students' high order thinking skills. *Journal for the Education of Gifted Young Scientists*, *7*(2), 237-251.

Wibowo, F. C., & Suhandi, A. (2013). Application of project-based science creative learning (SCL) model for improving cognitive learning outcomes and creative thinking skills. *Jurnal Pendidikan IPA Indonesia, 2*(1), 67-75.

Wilcox, B. R., Baily, M. C., Chasteen, S. V., Ryan, Q. X., & Pollock, S. J. (2015). Development and uses of upper-division conceptual assessments. *Physical Review Physics Education Research*, *11*(2), 020115.

Winarti, Cari, Suparmi, Sunarno, W., & Istiyono, E. (2017). Development of two tier test to assess conceptual understanding in heat and temperature. *Journal of Physics: Conference Series*, 795.

Yuliani, K., & Saragih, S. (2015). The development of learning devices based guided discovery model to improve understanding concept and critical thinking mathematically ability of students at islamic junior high school of Medan. *Journal of education and practice*, *6*(24), 116-128.

Yusrizal & Halim, A. (2017). the effect of the one-tier, two-tier, and three-tier diagnostic test toward the students' confidence and understanding toward the concepts of atomic nuclear. *Unnes Science Education Journal, 6*(2), 1-10.