Development of Assessment Instrument of Critical Thinking in Physics at Senior High School

T Sugiarti*, I Kaniawati and L Aviyanti

Post-graduate of physics education, Universitas Pendidikan Indonedia
Jl. Dr. Setiabudhi No. 229, Bandung 40154 Indonesia

*Corresponding author’s email : tiarsugiarti@student.upi.edu

Abstract. The result of preliminary study shows that the assessment of physics in school did not train students’ critical thinking skill. The assessment instrument just measured low cognitive aspects. Supposedly, critical thinking skill is trained in the assessment activity. The study aims to determine the characteristics and the quality of critical thinking skill instrument. It employs descriptive-qualitative method with research and development as the research design. The research participants are 35 students involved in the limited trial and 188 students in the wider trial from three public senior high school in Ciamis which in high level school. The data was collected through expert validation, tests and interviews. The results indicate that the characteristics of the assessment instrument of critical thinking skill is open-ended. The instrument fulfills some indicators namely analyzing argument, deduction, induction, and display information in the form of scenario, text, graphic and table. In addition, the data processing through V4 Anates program shows that the instrument reliability achieves 0.67 with high interpretation of 0.67 and the validity is 0.47 with enough interpretation. Thus, the assessment instrument of critical thinking skill in the form of open-ended essay meets the criteria of quality test, so it can use as instrument of assessment critical thinking skill.

1. Introduction
A preliminary study suggested that assessment instrument in Physics at school only measured the students’ basic thinking skill. Critical thinking skill, which includes to one of high level thinking skills, has not been examined by the teachers. It is due to the fact that they refer to the National Examination’s item sheets. The analysis of the National Examination’s item sheets in last five years, showed that 7% of them examines memory, 28% measures understanding, and 62.4% of the items evaluates application. Those three aspects include to basic thinking skills (Hawker: 2011).

Teachers said that they rarely analyzed each item given to the students because they do not have enough time to do it, so the teachers do not recognize the quality of each item they have made. According to Arikunto (2008), a test is considered good if it has two test requirements, namely validity and reliability. So, teachers have to analyze each item they employed.

Generally, teachers trained critical thinking skills during teaching by asking orally about physical phenomena. Students were expected to gain information and respond by stating opinion about the questions asked.

Critical thinking skills have not been examined in the form of item sheets because there are small amount of items measuring physics learning performance in terms of critical thinking skills. Whereas,
according to Tuncay (2006), assessment tools employed by teachers played important roles in assessing students’ achievement and development of students’ critical thinking skills. Therefore, assessment tools not only useful for examining the students’ mastery, but also to improve their thinking ability.

One of the alternatives that can be used to overcome the problems mentioned above is by developing assessment instrument of critical thinking skills in physics for senior high school. Thus, that development process showed characteristics and quality of instrument of critical thinking skills.

Critical thinking is defined as ability to give reasons (reasonable) and think reflectively focused on belief and behavior (Ennis: 2011). Ennis (2011) says that critical thinking skills include several aspects, namely basic explanation, foundation of making decision, conclusion, further explanation, supposition and combination, and other additional skills. In examining critical thinking skills, Ennis (2001) stated that open-ended problems are suitable to assess critical thinking because it can be adapted easily and more comprehensive. Open-ended problems are questions or problems that have more than one possible answer and more than one strategies to answer (Pelfrey, as cited in Azizah: 2013). Based on information complexity displayed on the problems, Ennis (2001) divided open-ended problems into three structures, namely high structure, medium structure, and minimal structure.

Stobaugh (2013) explains that in constructing an assessment instrument of critical thinking skills can be done by giving information on the problems. The information given are either in the form of scenario, concrete examples, or authentical tasks. In addition, they can be delivered by using illustrations (pictures, cartoons, advertisements), tables, graphics, and maps. The information can also be communicated by using quotations or texts.

2. Method

The present study employed a descriptive qualitative research method with research and development design. This research and development were conducted until the level of instrument development covering limited test and broader test. The participants of limited test were from one of the classes in a public senior high school in Kabupaten Ciamis which has 35 students. They were selected using purposive sampling. Whereas for broader test, 188 from those three public senior high school were involved.

The developed assessment instrument of critical thinking skills was examined on instrument validity sheets. Validity sheets contains assessments of instrument contents and constructions, and validated by three expert lecturers. Those assessment instruments were in the form of essays on open-ended problems, consist of six items, examined to the students to obtain empirical validity, so that the quality of developed instrument was ascertainable. Moreover, an interview was conducted to the participants in order to gain information about clarity of the developed problems. Consequently, those informations are helpful to improve the developing instrument of critical thinking skills.

Analysing items, including validity analysis, reliability, and discrimination power, were used to identify quality of instrument. Reliability on limited test was determined by a test-retest method, namely giving similar test to the same object in certain periods of time (Aikunto: 2008). Reliability was counted by employing Pearson’s correlation coefficient

$$r_{xy} = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{(N\sum X^2 - (\sum X)^2)(N\sum Y^2 - (\sum Y)^2)}}$$

Software Anantes V4 was employed to analyze the results of validity, reliability, discrimination power, and level of difficulty.

3. Finding and Discussion

There were nine items of assessment instrument which were validated by expert lecturers. They covered an aspect of basic explanation with indicator of argument analysis, and aspect of conclusion, including indicator of deduction and induction. The results of expert lecturer’s validation were shown in the table below:
Table 1. Recapitulation of the Results of Content Validation of Critical Thinking Skills Instrument in Physics at Senior High School

| Problem No. | CTS Indicators       | Problem Types | Lecturer 1 | Lecturer 2 | Lecturer 3 |
|-------------|----------------------|---------------|------------|------------|------------|
| 1           | Argument analysis    | HS            | VWI        | VWI        | VWI        |
| 2           | Deduction            | MS            | VWI        | NV         | VWI        |
| 3           | Induction            | MS            | VWI        | VWI        | VWI        |
| 4           | Deduction            | HS            | VWI        | NV         | VWI        |
| 5           | Argument analysis    | MS            | VWI        | VWI        | VWI        |
| 6           | Induction            | MS            | VWI        | VWI        | VWI        |
| 7           | Argument analysis    | MS            | VWI        | NV         | VWI        |
| 8           | Deduction            | MS            | VWI        | V          | VWI        |
| 9           | Induction            | HS            | VWI        | VWI        | VWI        |

Note:
- MS: Medium Structure
- VWI: Valid with improvement
- V: Valid
- NV: Not Valid
- HS: High Structure

Table 1 shows that six out of nine developed problems of critical thinking skills in physics at senior high school number 1, 3, 5, 6, 8, and 9 were considered valid with improvement, both in terms of constructing interrogative sentences, spelling, picture information, and scoring column. Meanwhile, the other 3 items, namely number 2, 4, and 7 were considered not valid because they did not suit to indicators of critical thinking skills. Those three items were not discarded but they were revised by adding clearer information, changing interrogative sentence into more appropriate form according to the lecturer’s suggestions, so that the problems could be used to measure the students’ critical thinking skills.

There were six problems used in instrument test namely number 1, 3, 4, 5, 8 dan 9 because those six problems have fulfilled the indicators of critical thinking skills, namely argument analysis, deduction and induction. Those six problems also represented two types of problems of critical thinking skills. Problems number 1, 5 and 9 represented high structure, while numbers 3, 4, 8 corresponds to medium structure.

3.1. The Results of Limited Test of Assessment Instrument of Critical Thinking Skills
The analysis of the limited test problems was conducted by using Anates V4, with validity score of 0.47, interpreted adequate. Meanwhile, reliability was determined by a test-retest method and counted by employing product moment correlation(Arikunto: 81), with realiability score of 0.82, considered very high. Table 2 shows the results of problems analysis of limited test.
After analyzing, the problems considered unfavorable or less than adequate were changed or repaired. Decision making towards those problems need repairing was done with several considerations about the result of analysis or item validity ($r$), level of difficulty ($P$), and discrimination power ($D$). According to Mulyatiningsih (2009), if two out of three criteria of good problems have been fulfilled or consistent, the problems could be employed. On the contrary, if two out of three criteria of good problems have not been fulfilled, the problems must be revised. In addition to problem analysis, revision was also done by considering the problems readability by the students through an interview, asking the clarity of information and questions on the problems.

3.2. The Result of Broader Test of Assessment Instrument of Critical Thinking Skills

The result of broader test problems analysis processed with Anates V4 was 0.47 for validity and 0.67 for reliability test as a whole, interpreted adequate and high. Table 3 shows the results of problems analysis with broader test.

| Problems No | CTS Aspects     | Item Validity | Discrimination Power | Level of Difficulty | Conclusion          |
|-------------|-----------------|---------------|----------------------|---------------------|---------------------|
|             | Score | Category | Score | Category | Score | Category |             |
| 1           | Argument analysis | 0.48     | Adequate | 0.22     | Adequate | 0.40    | Medium   | Accepted |
| 2           | Deduction       | 0.65     | High     | 0.51     | Fine    | 0.54    | Medium   | Accepted |
| 3           | Induction       | 0.70     | High     | 0.34     | Adequate | 0.42    | Medium   | Accepted |
| 4           | Argument analysis | 0.58     | Adequate | 0.15     | Poor    | 0.39    | Medium   | Accepted with revision |
| 5           | Deduction       | 0.75     | High     | 0.30     | Adequate | 0.29    | Medium   | Accepted with revision |
| 6           | Induction       | 0.53     | Adequate | 0.29     | Adequate | 0.44    | Medium   | Accepted |

After analyzing, the problems considered unfavorable or less than adequate were changed or repaired. Decision making towards those problems need repairing was done with several considerations about the result of analysis or item validity ($r$), level of difficulty ($P$), and discrimination power ($D$). According to Mulyatiningsih (2009), if two out of three criteria of good problems have been fulfilled or consistent, the problems could be employed. On the contrary, if two out of three criteria of good problems have not been fulfilled, the problems must be revised. In addition to problem analysis, revision was also done by considering the problems readability by the students through an interview, asking the clarity of information and questions on the problems.

3.2. The Result of Broader Test of Assessment Instrument of Critical Thinking Skills

The result of broader test problems analysis processed with Anates V4 was 0.47 for validity and 0.67 for reliability test as a whole, interpreted adequate and high. Table 3 shows the results of problems analysis with broader test.
After analyzing problems, decision making towards those problems need repairing was done with several considerations about the result of analysis or item validity (r), level of difficulty (P), and discrimination power (D). In addition to problem analysis, revision was also done by considering the problems readability by the students through an interview, description of students’ answer, and analysis of problems compatibility to physics for senior high school. The followings an examples of problems analysis of assessment instrument

1) Problems number 3

Problems number 3 covers the aspects of conclusion, indicating induction of relationship between rising of temperature and heat based on the results of experiment. This problem is a kind of medium structure and outlined in the form of table of experiment data containing information about heat, initial temperature, final temperature, and temperature changing. In this problem the students were asked to induce by generalizing based on the existing data and completing their answer with graphic.

An analysis to students’ answer showed that they were able to conclude the relationship between heat and rising of temperature. They could make a conclusion based on systematic data. The students were also able to conclude that temperature changing was influenced by number of heat given. They could also convert the information into the form of graphic.

Based on the interview, the clarity of problems number 3 was clear enough, both information given and the questions asked. Problems number three has suited the materials given in the classroom. The students have learned about some factors influencing amount of heat. This problem has also suited the quality of good problems based on an analysis conducted and displayed on table 3. Thus, problems number 3 could be used as assessment instrument of critical thinking skills without any revision.

3.3. Characteristics of Assessment Instrument of Critical Thinking Skills in Physics at Senior High School

The problems analysis with broader test showed that those problems have validity score of 0.47, interpreted “adequate” and reliability score of 0.67, interpreted “high”. It resembled Amalia (2012) that the characteristics of problems to examine critical thinking skills are those which have validity and reliability scores of more than adequate.

The characteristics of problems to examine critical thinking skills are those which written in the form of open-ended problems divided based on complexity of information, namely high structure, medium structure, and minimal structure (Ennis: 2001). The problems which were developed in this research were high structure and medium open-ended essay problems. High structure problems give information in form of text to the students. The high structure problems to examine critical thinking skills was provided with three indicators, namely argument analysis (number 1), deducting (number 5), and inducting (number 6). Problems with medium structure were examined by giving simpler information than that of high structure.

| Problems No | CTS Aspects         | Item validity | Discrimination power | Level of Difficulty | Conclusion       |
|-------------|---------------------|---------------|----------------------|---------------------|------------------|
|             |                     | Score         | Score                | Score              |                  |
|             |                     | Category      | Category             | Category           |                  |
| 1           | Argument analysis   | 0.55          | Adequate             | 0.34               | Adequate         | Medium           | Accepted         |
| 2           | Deduction           | 0.65          | High                 | 0.47               | Fine             | Medium           | Accepted         |
| 3           | Induction           | 0.66          | High                 | 0.40               | Adequate         | Medium           | Accepted         |
| 4           | Argument analysis   | 0.29          | Low                  | 0.07               | Poor             | Medium           | Accepted         |
| 5           | Deduction           | 0.82          | Very high            | 0.55               | Fine             | Medium           | Accepted with revision |
| 6           | Induction           | 0.44          | Adequate             | 0.16               | Poor             | Medium           | Accepted         |
Such problems were examined through problems number 2 (indicator of deduction), number 3 (indicator of induction), and number 4 (argument analysis).

Moreover, this assessment instrument of critical thinking skills were also ordered based on two aspects of critical thinking skills, namely involving basic explanation and concluding. Those two aspects covers three indicators, namely argument analysis, deducting, and inducting. The developed aspects have suited to aspects of critical thinking skills stated by Ennis (2011), namely involving basic clarification, foundation of decision making, inferring, explaining further, approximating and combining, and additional skills. Aspects and indicators of critical thinking skills measured in this study were explained below.

1) Involving basic clarification
The aspects of Involve basic clarification were examined through problems number 1 and 4. This aspect presented an ability to analyze arguments, which could be tested by identifying irrelevancy and considering possible answer as well as searching for similarities and differences between something.

2) Inferring
The aspect of inferring presented someone’s ability to infer both in inductive and deductive way. Indicator of deducting was examined to question number two and five. Those problems were presented by giving some statements which are assumed correct, alternatives that consists of likely correct conclusion and asking students to check those conclusions.

Indicator of induction was examined through problems number 2 and 6. Those problems were given by presenting data. The students were asked to make a generalization based on information given on the problems. Critical thinking skills in the present study adopted those according to Ennis (2011). One of the characteristics of this type of problems is that there are some relevant information so that the students could think reasonably and give reasons to make decision of what should be believed and done. In this case, information were important things to construct the problems of critical thinking skills.

Based on the description of problems that have been explained before, table 4 following displayed the characteristics of assessment instrument of critical thinking according to information attached to the problems.

| Problems No | CTS Aspects       | CTS Indicators | Description of Information attached to the problems |
|-------------|-------------------|----------------|---------------------------------------------------|
| 1           | Involving basic clarification | Argument analysis | Text about phase transition                         |
|             |                   |                | Graphic about phase transition of CO2              |
|             |                   |                | Graphic about phase transition of water            |
| 2           | Concluding        | Deducting      | Picture showing percentage of sun radiation        |
|             |                   |                | Conclusions made based on the picture              |
| 3           | Concluding        | Inducting      | Table showing results of experiment to determine relationship between heat and change of temperature |
| 4           | Involving basic clarification | Argument analysis | Scenario containing opinions of two experts about AC installation |
| 5           | Concluding        | Deduction      | Text containing information about ice field in Antarctic |
|             |                   |                | Statements showing conclusion of the text          |
| 6           | Concluding        | Inducting      | Scenario of daily life                             |
|             |                   |                | Table showing amount of metal’s conductivity       |
|             |                   |                | Table showing elements diversity in the universe  |
Table 4 showed that information attached to the problems are in the form of graphic, text, table, and scenario. Such information suited to those used to test critical thinking skills employed by Stobaugh (2013), namely scenario, real examples, and authentic assignments. Visual displays were in the form of illustration including advertisement, painting, picture and cartoon, diagram and table, map and graphic. Other information were in the form of quotations which consists of short quotation or text. Those informations were combined in a problem, for instance scenario and picture or scenario and table. These studies found a characteristic that information displayed in the instrument were scenario, visual information in the form of table and graphic as well as text.

An experiment in the form of results of lecturer’s judgement and interviews to the students that have been elaborated before, showed a characteristic of assessment instrument of critical thinking skills: the problems should be written in effective sentence, suited to the standard Bahasa Indonesia. The information should be clear. For example, the picture should be colored and the paragraph should be written simply. Information content should be related to materials learned by the students. If it was a new context, it would better to asked questions related to materials that have been learned by the students. It was to avoid who could not answer the questions. Assessment was aimed at measuring and examining purpose of learning (Surapranata: 2007).

4. Conclusions
Present study concluded that the characteristics of assessment instrument of critical thinking skills in physics for senior high school covering two aspects, namely involving basic clarification indicating argument analysis, and the aspects of inferring indicating deduction and induction. As a result, further studies, discussing instrument development based on other aspects and indicators of critical thinking skills, are needed either refers to the same or different expert.

The problems developed in this study are high structure open-ended essays which have complex enough information, and medium structure with simpler information. Therefore, supplementary studies should be conducted to explain about assessment instrument of critical thinking skills in different form of problems, like multiple choice.

Assessment instrument of critical thinking skills is a kind of problems that contains various informations. The information to use is scenario, table, graphic, and text. Thus, further studies about the development of assessment instrument of critical thinking skills are required, based on the types of given information, such as authentic assignment, comic, and quotation.

Based on problems analysis, assessment instruments developed in this study have fulfilled the quality of a good test by validity and reability scores of 0.47, interpreted adequate, and 0.67, interpreted high. However, according to the results of interview, one out of six problems were revised by changing the content of the text, so that the problems suited the materials learned by the students. Two problems were changed in visual display, and another one was revised by reducing given information.

Problems development in this study only covers heat. Therefore, future studies were required to develop assessment instrument of critical thinking skills in other topics learned in senior high school. In addition, the problems should be written by considering materials that have been learned by the students.

5. References
[1] Amalia S N 2012 Pengembangan Soal Esai Berpikir Kritis dan Profil Pencapaianya di SMA Negeri Kota Bandung Pada Tema Sistem Organ. A Thesis (School of Postgraduate Studies UPI Bandung: Unpublished)
[2] Arikunto S 2008 Dasar Dasar Evaluasi Pendidikan (Jakarta: Bumi Aksara)
[3] Azizah Y N 2013 Pengembangan dan Analisis Soal Open-Ended Problem Dalam Mengukur Keterampilan Berpikir Kreatif pada Materi Minyak Bumi (An Undergraduate Thesis UPI Bandung: Unpublished)
[4] Ennis R H 2001 *Critical Thinking Assessment* [Online]
[5] Ennis R H 2011 *The Nature of Critical Thinking: An Outline Of Critical Thinking Disposition and Abilities* [Online]
[6] Hawker B 2011 *How To Assess High-Order Thinking In Your Classroom* [Online]
[7] Mulyatiningsih E 2012 *Metode Penelitian Terapan Bidang Pendidikan* (Bandung: Alfa Beta)