Domain-specific critical thinking skills: Designing and analysis the instrument test on thermochemistry

E Nursa’adah¹*, D Sari² and I Affifah²

¹Science Education Department, Universitas Bengkulu, Jl. W.R Supratman Kandang Limun Bengkulu 38371 A, Indonesia
²Chemistry Education Department, Universitas Sultan Ageng Tirtayasa Banten, Indonesia

*euis@unib.ac.id

Abstract. Chemistry learning needs an assessment to evaluate ways of students’ thinking. One them is critical thinking skill test. The purpose of this study is designing domain-specific critical thinking (DS-CT) skill on thermochemistry concept. There are 29 DS-CT Skills test to measure students’ DS-CT skills. Content Validity Ratio (CVR) and Confirmatory Factor Analysis (CFA) were evaluated to analysis the validity of DS-CT Skills test. Cronbach’s Alpha evaluated to analysis the reliability. Research and Development was used as research design. The participants of this study are 185 senior high school students in Cilegon-Banten-Indonesia. Data are obtained from expertise judgment, empirical test, and interview. The results showed that validity of DS-CT skills test on thermochemistry concept based on CVR and CFA is 0.723 (good criteria). Cronbach’s Alpha test showed 0.652 (high criteria). It can be conclude that this DS-CT Skills test is feasible to measure students’ DS-CT skills on thermochemistry domain knowledge. Students’ CT skill on domain thermochemistry are 46% on present and provide arguments, 27% make and assess claims, and 21% analysis arguments.

1. Introduction
The rapid development of science and technology not only provides benefits for human survival but also raises various complex and multidimensional problems. The problems require a person to have good endurance and thinking skills in analyzing, understanding and the ability to solve problems well [1]. In the educational area, Partnership for 21st Century Skills [2] states that innovation skills, technology & information, survival skills are needed by students. Wagner [3] emphasizes eight important survival skills in the 21st century era, namely: 1) communication skills; 2) critical and creative thinking; 3) reasoning skills; 4) Ennis [4] provides a definition of interpersonal skills; 5) multicultural/multilingual literacy; 6) problem solving; 7) information/digital literacy; and 8) technological skills. These skills should be possessed by Indonesian students in order to be able to compete at the international level.

International studies already discussed 21st-century skills such as Trends in International Mathematics and Science Study (TIMSS) [5] showed that Indonesia only achieved 5% on reasoning issues as one of higher order thinking skills. Meanwhile, one of the criteria for a person's success in life in this era is determined by thinking skills in order to be able to solve the problems from simple to complex problem in their daily life. Science education has an important role in increasing these skills in accordance with the objective of science education.
The development of high order thinking skills at the school level has been done by researchers and teachers. One of the skills is critical thinking skills. Students at senior high school level are able to think abstractly and complexly and it makes them able to solve the problem. That is the assumption that critical thinking skills can be applied at the high school level. Critical thinking skill itself is a well-organized mental process that plays a role in the process of making decisions to solve problems by analyzing and interpreting data in scientific inquiry activities [6]. It means that students without critical thinking skills make them not able to analysis and managing strategies in problem-solving [7]. Applying critical thinking in a domain specific knowledge it called domain specific critical thinking skills (DS-CT skills) [8].

CT skills applied in learning are supported by the assumption that students generally can achieve it through practice so their thinking skills always develop, can be taught, and also be learned [9]. The implication of these assumptions is to create an evaluation system that can renew the mindset of students. To understand students 'success in developing CT skills, the instruments that can measure students' CT skills are needed.

Some CT skills instruments that have been developed are multiple-choice forms. Meanwhile, Stiggins [10] and Ku [11] states that to measure CT skills that should not use multiple-choice tests. Open-ended questions allow students to answer problems in many ways and many answers. It provide students to describe facts, logical arguments to consider and choose a source to answer or solve the problems. This statement is supported by Lai's which states that open-ended questions can be used to measure CT skills. The questions arranged not only make students to recall information but also require students to use information in new contexts. In addition, Ennis [4] suggests that assessing CT skills should use open-ended questions because it is easier to adapted and more comprehensive.

CT skills test has different characteristics compared to the common tests. According to Gokhale [12], a CT skill test is a question that involves the analysis, synthesis, and evaluation of a concept. In applying CT skills indicators to the item questions, students are encouraged to: 1) determine the consequences of a decision; 2) Identify the assumptions; 3) formulate the main problem; 4) find bias based on different points of view; 5) Reveal the cause of a phenomena; 6) Selecting factors that support a decision [13].

Chemistry is a part of science that applies higher-order thinking skills in every concept and plays an important role in developing students' CT skills [14]. This is easily related to the phenomenon of daily life and its concept. One of the chemistry concepts is thermochemistry. For instance, the process of fat metabolism in the body. There are many arguments needed to explain the process. Students are expected to have CT skills-analyzing arguments - in identifying arguments to choose supporting or opposing the arguments given.

In this study, the CT skill indicators developed include categorizing, analyzing arguments, evaluating claims, determining alternatives, and presenting arguments in thermochemistry concepts [15]. Based on the background above, the objective of this research are design valid and reliable DS-CT skills on thermochemistry concept and analyze students CT skills.

2. Methods

Research and Development (R&D) is used as a research design, which aims to develop valid and reliable CT skills instruments on thermochemistry concepts. There are three general steps in this research design consist of design, test, and revise [16] research design is presented in Figure 1. Valid and reliable CT skills test is one of research product. In addition, this research also used descriptive research method to examine students' CT skills.

![Figure 1. Research design.](image-url)
Research subjects were 185 senior high school students from several schools in one of city in Banten Province Indonesia. The research procedure is performed as follows: 1) Designing the CT skills test, the researcher determines the purpose of the test, competencies, thermochemistry concepts and scoring guidelines; 2) CT skills test is validated by 15 expertise in chemistry and chemistry education. Content validity ratio (CVR) is determined based on expertise input. CT skills tests were tested on 150 students in several senior high schools, then confirmatory factor analysis (CFA) is tested. Cronbach alpha is used to analyze reliability test; 3) Revises CT skills test and data analyze; 4) analyze of students’ CT skills on thermochemistry concept. The research procedure can be seen in Figure 2.

Figure 2. Research procedure.

The instrument in this study was a DS-CT skill test on the thermochemistry concept. The instrument was validated by 15 expertise in chemistry and chemistry education. Validated aspects are the appropriateness of the questions used with CT skill indicators, the appropriateness of the key answers and the questions, the scoring rubric, and good language. The instrument contains CT skills indicators
proposed by Facione [15], they are: 1) categorization, 2) analysing arguments, 3) assessing claims, 4) determining alternatives, 5) presenting arguments.

Expertise validity is analyzed using Content Validity Ratio (CVR) [17]. Ratio criteria presented in Table 1. If calculated CVR greater than CVR table, the item declared as valid item.

| Numbers of Panelists | Minimum Acceptable CVR Value |
|----------------------|-----------------------------|
| 5                    | 0,99                        |
| 6                    | 0,99                        |
| 7                    | 0,99                        |
| 8                    | 0,75                        |
| 9                    | 0,78                        |
| 10                   | 0,62                        |
| 11                   | 0,59                        |
| 12                   | 0,56                        |
| 13                   | 0,54                        |
| 14                   | 0,51                        |
| 15                   | 0,49                        |
| 20                   | 0,42                        |
| 25                   | 0,37                        |
| 30                   | 0,33                        |
| 35                   | 0,31                        |

The authors also performed a confirmatory factor analysis (CFA) to test accurately and consistently conceived indicators and what indicators act as the constructs studied.

3. Results and discussion

3.1. Validity and reliability of critical thinking skill test
There are 29 questions to measure DS-CT Skills on thermochemistry domain. Each questions has CVR more than 0,49 (Valid). Therefore, there is one item question that has value of CVR less than 0,49 (Unvalid). All CVR for each question from 15 panellists can be seen at Figure 3.

![Figure 3. CVR Results from 15 panelists for each questions.](image-url)
The results of expert validation include several aspects, they are content, language, and construction. Qualitatively, the results of expert validation show that there are some questions that need revision in readability, focus questions, add the ΔH symbol to the chemical reaction equation, and DS-CT Skills indicators.

Confirmatory Factor Analysis (CFA) is used to test whether the number of factors obtained empirically matches the number of factors that have been theoretically arranged or test hypotheses about the existence of constructs. The KMO and Barlett's tests is .689.

The next analysis is to examine the Measure of Sampling Adequacy (MSA). This aims to determine whether the sampling process is adequate or not, which can be seen from the Anti-Image Correlation (>0.5a). In the Anti-Image Matrics obtained 3 (three) questions that have MSA values less than 0.5, they are questions numbers 2, 6, and 22. Therefore, to be able to proceed to the next stage of the analysis of the questions number 2, 6, and 22 excluded from analysis. The distribution of problem indicators is presented in Table 2.

| Distribution of Questions | Factor | Name of Factor     | Thermochemistry Concept                                      |
|---------------------------|--------|--------------------|---------------------------------------------------------------|
| 2, 6, 1, 22               | 1      | Categorization     | Aspects of Thermochemistry, systems, exothermic, and enthalpy of combustion |
| 8, 9, 13, 14, 10, 15, 16, 24, 11 | 2 | Analysis of Arguments | Exothermic, enthalpy of reaction, thermochemistry equation, enthalpy of combustion, calorimeter, and measurement enthalpy |
| 29, 27, 25, 21, 20, 18, 19, 17, 23 | 3 | Make a claim       | Calorimeter, enthalpy, Hess Law, standard enthalpy of bonding, and application |
| 4, 26, 7, 28, 5, 3         | 4 | Decision of Alternative | Endothermic, exothermic, standard enthalpy of bonding |
| 12                        | 5 | Provide an Argument | Thermochemistry equation                                    |

Second test was conducted due to invalid variables in the previous test. All three questions were excluded from the analysis and re-tested for valid variables. The results showed a strong significance for Bartlett test (Chi-Square = 1562.749 and p < 0.001). In addition Kaiser-Meyer-Olkin (KMO) was measured to be .723. It concludes that 72.3% of the variance can be explained by these factors. MSA examination continues. The final results of this second test are presented in Table 3. Cronbach Alpha test show that reliability of test is 0.652. This shows that the 26 item items have a high reliability value.

| Distribution of Questions | of Factor | Name of Factor   | Thermochemistry Concept                                      |
|---------------------------|-----------|------------------|---------------------------------------------------------------|
| 29, 27, 25, 21, 20, 18, 19, 17, 23 | 1 | Make claims     | Calorimeter, enthalpy, Hess law, standard enthalpy of bonding, and application |
| 8, 9, 13, 14, 10, 15, 16, 24, 11, 12, 1 | 2 | Analysis of Arguments | Thermochemistry aspect, exothermic, thermochemistry equation, and enthalpy of combustion |
| 4, 26, 7, 28, 5, 3          | 3 | Provide an Arguments | Endothermic, exothermic, and standard enthalpy of bonding |
3.2. Students’ DS-CT skills on thermochemistry domain

Students’ DS-CT Skills on each indicators is presented in Figure 4.

![Figure 4. Percentage of students’ DS-CT skills.](image)

The first indicator is making and assessing claims, this indicator requires students to have skills in assessing the credibility of statements or other representations that are intended to determine the statement given is true or false based on known knowledge. For example in question number 19, this question requires students to assess the credibility of the statement regarding enthalpy and include the reason. 52% from 185 students are able to answer. Most of them only write down statement points that are considered true, without including supporting reasons or explanations. Therefore, the achievement of students’ CT skills on this indicator is 27%.

The second indicator is analysis arguments; this indicator requires students to be able to identify the assumptions implicit in reasoning. It aims to support or oppose some claims, opinions or one's point of view. For example in problem number 11, this question requires students to identify the argument “250 grams of fat can meet energy needs for one day”. 17% of the 185 students who answered, most of them just rewrote the thermochemical equation contained in the problem; there were some students who answered right. Unfortunately, they were no complete with the right calculation. Therefore, the achievement of students’ CT skills on the indicator analysing argument is only 21%.

The third indicator is to present and provide an argument, this indicator requires students to identify and express evidence and counter-evidence about the problem being asked. For example in problem number 3, this question requires students to analyse two chemical reaction processes, then students are asked to provide arguments accompanied by evidence about the question why the two reactions belong to the endothermic reaction. 96% of the 185 students who answered, most of them can provide arguments accompanied by evidence of the reasons why the two reactions are classified as endothermic reactions. Therefore, the achievement of students’ CT skills on this indicator is 46%.

Figure 4 showed that most of students have a low ability level on two from the three indicators of DS-CT Skills. Indicators that have not been achieved by students are assessing claims and analysis arguments. Students are not accustomed to working on CT skills that are required to use their skills in interpreting, analysing, and evaluating. Thinking skills are not an instant learning outcome that can be directly measured by two to three times learning, then declared good or not good. Based on the results of research from Richmond [18], states that continuous process and practice is needed to be able to change students’ thinking skills.

4. Conclusion

DS-CT skills instrument on thermochemistry developed in this study includes three indicators of critical thinking skills adopted from Facione. There were 26 valid and reliable Questions based on the results of the CVR, CFA, and reliable tests. The validity of the instrument was 0.723 (good category) and the reliability of the instrument with a Cronbach's Alpha value is 0.652 (high). The percentage of students’
DS-CT skills attainment is 27% on indicators make claims, 21% on indicators analysing arguments, and 46% on indicators provide arguments.

Acknowledgments

Many thanks are due to all students who volunteered to participate in the study.

References

[1] Kartimi and Liliasari 2012 J. Pendidikan IPA Indonesia 1(I) 21-26
[2] Partnership for 21st Century Skills 2013 [Online] Retrieved from: www.21stcenturyskills.org
[3] Wagner T 2008 The Global Achievement Gap (New York: Basic Books)
[4] Ennis R H 1993 Critical Thinking Assessment (UK: Theory into Practice)
[5] TIMSS 2011 International Results In Science (Boston: The TIMSS & PIRLS International Study Center, Boston College)
[6] Johnson E B 2002 Contextual teaching and learning: What it is and why it's here to stay (California, US: Corwin Press)
[7] Axviarani V N and Widodo A T 2014 Pengembangan Perangkat Pembelajaran Kesetimbangan Kimia Dengan Strategi Probex Berbasis Keterampilan Berpikir Kritis Chemistry in Education 3(II)
[8] McPeck J E 1990 Educational Researcher 19(4) 10-12
[9] Nickerson R 1988 Review of Research in Education 15(3) 3-57
[10] Stiggins R J M 1994 Student Centered Classroom Assessment (New York: McMillar College Publishing Company)
[11] Ku K Y L 2009 Assessing student’s critical thinking performance: urging for measurement using multi-response format Thinking Skills and Creativity 4 70-76
[12] Gokhale A A 1995 J of Tech. Edu 7(1) 22-30
[13] Langrehr J 2006 Thinking Skills Edisi Pertama (Jakarta: PT. Gramedia)
[14] Zhou Q, Huang Q and Tian H 2013 Developing students’ critical thinking skills by task-based learning in chemistry experiment teaching Creative Education 4(XIIA) 40-45
[15] Facione P A 1990 Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction (ERIC: California State University, Fullerton)
[16] Wilson M 2005 Constructing measures: An item response modelling approach (Mahwah: Lawrence Erlbaum Associates, Inc. Publishers)
[17] Wilson F R, Pan W and Schumsky D A 2012 Measurement and Evaluation in Counseling and Development 45(3) 197-210
[18] Richmond J E D 2007 J Int. Edu. 8(I) 1-29