DEVELOPMENT OF INITIAL ABILITY TEST INSTRUMENT BASED ON SCIENCE GENERIC SKILLS

Arta Elsa Putri Ambarita 1*, Ratu Evina Dibyantini 2, Ani Sutiani 3, Zainuddin Muchtar 4, Marudut Sinaga 5

1,2,3,4,5 Chemistry Department, Chemistry Education Study Program, Faculty of Mathematics and Natural Sciences, Universitas Negeri Medan
Jl. Willem Iskandar / Pasar V, Medan 20221, Indonesia
*artaelsaabr@gmail.com

ABSTRACT

This study aims to develop an initial ability test instrument based on science generic skills in the form of a scholastic test and develop a final chemistry evaluation test instrument in the topic of thermochemistry as well as find out the validity of the content and construction validity of the instrument. Research and development research, with a 4D development model consisting of 4 main stages, namely: Define, Design, Develop and Disseminate, due to time constraints this research was only up to develop. The results of the content validity test by experts obtained an average percentage of material aspects, namely 87.08%, construction aspects, namely 90.41%, language aspects, 92.5% and additional rules, 96.67%, so the question instrument is suitable for use. Based on the construction validation test data, it was found that of the 20 questions on the scholastic potential test tested, there were 19 questions that were valid and considered appropriate, while 1 other item was invalid, while of the 20 questions on the chemistry test tested, there were 17 questions that were valid and considered feasible, leaving 3 questions other invalid. For further research, it is hoped that this instrument can be implemented.
Introduction

In the 21st century, the development of technology and information is running very fast, the culture is starting to develop and the development of education at the international level makes this an external challenge that must be faced, it is necessary to improve the quality of education to produce qualified and capable Human Resources (HR), compete in the era of globalization.

**PISA (the programme for international student assessment)** is a program to measure achievement in the fields of math, science and reading literacy. The current PISA assessment has been used as a reference and evaluation of the quality of education of a participating country from PISA. Based on the 2018 PISA results in the reading ability category, Indonesia is ranked 74th out of 79 countries, while for the assessment of math skills and science skills, Indonesia is ranked 73rd and 71st out of 79 participating countries in PISA. The achievement of Indonesia’s ranking in the PISA assessment has always been constant since the beginning of Indonesia’s participation in the assessment, namely from 2000 to 2018.

With consistent results being ranked lower brings the consequence of thinking that the quality of Indonesia’s education is not in accordance with global community standards and is below other countries, in the world (Hewi & Shaleh, 2020). The existence of a gap between expectations and reality in adapting to world changes and developments in the era of globalization is a problem that must be overcome, especially in the world of education. Therefore it is necessary to develop education that can optimally increase the potential and abilities of students, so that they are able to adapt to circumstances and changes that occur and are able to work collaboratively in solving life’s problems (Sutiani et al., 2022).

Ausubel's theory states that in helping students inculcate new knowledge from a material, it is very necessary to have an initial concept that students already have that relates to the concept to be studied so that it determines the success or failure of a learning process, therefore students' initial abilities are very important in learning (Rasvani & Wulandari, 2021). Ausubel's meaningful learning theory emphasizes linking old knowledge with new knowledge into students' cognitive structures, Ausubel's theory can help students find the meaning of learning (Kinasih & Sinaga, 2020).

Chemistry is part of the natural sciences which has abstract characteristics. Chemistry is the science that studies the structure, composition, properties and changes in matter and the energy that accompanies these changes. Chemistry can also make an important and meaningful contribution to the development of applied sciences, such as agriculture, health and fisheries and technology(Putri et al., 2021).

Thermochemistry is one of the subject matters of chemistry class XI IPA. Thermochemistry is a mathematical subject matter and requires a strong understanding of concepts that is often considered difficult by students. Students tend to memorize thermochemical material without understanding the concept, thus causing students to forget more easily, besides that thermochemistry is considered difficult by students because there are many chemical equations and calculations (Warih et al., 2015).

Students’ initial abilities are abilities that students already have before they take part in the learning that will be given. The initial ability of students is important for the teacher to know before he starts with his learning, because then it can be known; does the student already have or knowledge which is a prerequisite to follow the learning: The extent to which students already know what material will be presented. By knowing these two things, the teacher will be able to design learning better (Razak, 2018), this is in accordance with Ausubel’s theory which states that the learning process will go well if new subject matter or information can adapt to the cognitive structure that someone already has (Suryani et al., 2020).

By paying attention to students’ initial abilities, the teacher can determine the portion of the material for students in order to develop cognitive processes so that students can improve their high-order thinking skills well. Students with high initial abilities are better able to develop their cognitive processes so that it is easier to do higher-level thinking questions compared to students with low initial abilities. Scholastic tests in the form of verbal reasoning and numerical reasoning questions can be used as instruments to measure students' prior knowledge.

Scholastic ability is a combination of verbal and numerical thinking abilities, the combination of both scores being a good predictor for completing assignments in between academic
Development of Initial Ability Test Instrument based on Science Generic Skills

To determine the achievement of learning outcomes, it is necessary to evaluate learning outcomes. Evaluation of learning outcomes aims to determine the development and progress of student learning outcomes after the learning program, as well as to collect data and information in an effort to improve learning activities to achieve the goals set by the curriculum.(Badu, 2013). Learning evaluation cannot be separated from learning objectives. Without learning objectives, it is not possible to assess the extent to which students achieve in learning, because each assessment requires a criterion that becomes a reference in assessing the object. Learning objectives are the main criteria in learning evaluation.

Based on the results of research conducted on class XI high school students at a public high school in Bandung, it was found that students' generic science skills for indicators of constructing concepts, logical frames, indirect observation and causal law were classified as low and modeling was classified as moderate (Agustin, 2014), which is in line with research Ermanawi et al., (2019) which states that students' generic science skills are still relatively low. Based on research on 5 grade XIIPA SMAN I Lawang academic year 2012-2013 it is known that students with high initial ability tend to have high final test scores, while students with low initial ability tend to have low final test scores (Irawati Kartika, 2014). Students' initial abilities and generic science skills have an influence on improving students' cognitive learning outcomes (Muthmainah et al., 2019). The low initial ability of students results in low student interest in learning, difficulty understanding the material, and a lack of students' ability to solve problems. end of chemistry as well as knowing the material, and a lack of students' ability to solve problems.

Science generic skills are the basic (generic) abilities needed to train students' scientific work so that they can produce students who are able to understand concepts, solve problems, and other scientific activities, and are able to learn independently effectively and efficiently (Maulina et al., 2017). This skill is an intellectual ability resulting from a combination or complex interaction between scientific knowledge and skills (Dibyantini & Azaria, 2020). To determine the achievement of learning outcomes, it is necessary to evaluate learning outcomes. Evaluation of learning outcomes is to determine the development and progress of student learning outcomes after completing the learning program, as well as to collect data and information in an effort to improve learning activities to achieve the goals set by the curriculum.

Science generic skills provide opportunities for students to be actively involved in learning so that interactions occur between skills and concepts, principles and theories that have been discovered or developed.(Rosidah et al., 2017). Generic skills are cognitive strategies that can be related to cognitive, affective, and psychomotor aspects that can be learned and left behind in students (Ulia, 2017). By mastering generic science skills, students can learn and understand chemical concepts more easily(Izetbigovic et al., 2019). There are eight KGS indicators in science learning, namely: (1) Direct and indirect observation, (2) Awareness of scale, (3) Symbolic language, (4) A logical framework according to principles, (5) Logical inference, (6) Law causation, (7) Mathematical modeling, and (8) Building concepts (Rosidah et al., 2017).

Numerical ability is intelligence in using numbers and reasoning (logic) including in the areas of calculation, classifying and categorizing information, thinking with abstract concepts to find relationships between one thing and another, numerical ability is ability related to accuracy and speed in use basic arithmetic functions. Verbal ability is the ability to explain thoughts or the ability to link various information obtained and make hypotheses, with verbal abilities students are expected to be able to explain their thoughts and relate the information they obtain and be able to make a hypothesis (Irawan & Kencanawaty, 2017).

Science generic skills are the basic (generic) abilities needed to train students' scientific work so that they can produce students who are able to understand concepts, solve problems, and other scientific activities, and are able to learn independently effectively and efficiently (Maulina et al., 2017). This skill is an intellectual ability resulting from a combination or complex interaction between scientific knowledge and skills (Dibyantini & Azaria, 2020). To determine the achievement of learning outcomes, it is necessary to evaluate learning outcomes. Evaluation of learning outcomes is to determine the development and progress of student learning outcomes after completing the learning program, as well as to collect data and information in an effort to improve learning activities to achieve the goals set by the curriculum.

Science generic skills provide opportunities for students to be actively involved in learning so that interactions occur between skills and concepts, principles and theories that have been discovered or developed.(Rosidah et al., 2017). Generic skills are cognitive strategies that can be related to cognitive, affective, and psychomotor aspects that can be learned and left behind in students (Ulia, 2017). By mastering generic science skills, students can learn and understand chemical concepts more easily(Izetbigovic et al., 2019). There are eight KGS indicators in science learning, namely: (1) Direct and indirect observation, (2) Awareness of scale, (3) Symbolic language, (4) A logical framework according to principles, (5) Logical inference, (6) Law causation, (7) Mathematical modeling, and (8) Building concepts (Rosidah et al., 2017).

Research Methods

This study uses research and development methods Research and Development. The research was conducted using the 4D development model which consisted of 4 stages namely; Define, Design, Develop, and Disseminate. The reason researchers chose this model is because the 4D model's stages are programmed, simple, easy to understand and its implementation is more systematic.

Defining Stage
This stage is the needs analysis stage, in product development the developer needs to refer to the development requirements, analyze and collect information on the extent to which development needs to be carried out. This stage includes 2 main steps, namely: curriculum analysis and analysis of the formulation of learning objectives.

Planning Stage

The preparation of test standards is a step that connects the defining stage with the design stage. The preparation of test standards is based on the results of the analysis of the specifications of learning objectives and student analysis. From this, a grid of learning outcomes tests is compiled. The test is adjusted to the students' cognitive abilities and the scoring of the test results uses an evaluation guide that contains scoring guides and key answer questions. This stage consists of 3 steps, namely: (1) Determination of the form of the instrument, (2) Preparation of the grid, and (3) Design of the instrument.

Development Stage

The development stage is the stage to produce a development product. This stage includes: device validation by validators followed by revisions, then limited trials are carried out with students (Maydiantoro, 2021).

In this study, data were obtained in the form of qualitative and quantitative data. Qualitative data were obtained from assessments and suggestions for improvement from expert respondents (2 lecturers and 4 teachers) in the form of content validity data on initial ability test instruments sharing generic science skills, while quantitative data were obtained from the results of instrument trials which had previously been validated by expert validators. The test instrument was tested on one class, namely XI D of SMA Negeri 2 Percut Sei Tuan, with a total of 36 students to determine the validity of the construction. The test instrument was composed of 20 scholastic test questions and 20 chemistry learning outcomes evaluation questions, each of which had 5 answer choices (A, B, C, D, and E). The researcher gave test instrument sheets and work answer sheets then gave 30 minutes each to work on the scholastic test questions and the chemistry learning outcomes evaluation test. Small class tests were conducted to determine validity, reliability, level of difficulty, discriminating power and distractors.

Result and Discussion

Findings of Define Stage

This stage is the needs analysis stage, in product development the developer needs to refer to the development requirements, analyze and collect information on the extent to which development needs to be carried out. This stage includes 2 main steps, namely: curriculum analysis and analysis of the formulation of learning objectives.

Planning Stage

The preparation of test standards is a step that connects the defining stage with the design stage. The preparation of test standards is based on the results of the analysis of the specifications of learning objectives and student analysis. From this, a grid of learning outcomes tests is compiled. The test is adjusted to the students' cognitive abilities and the scoring of the test results uses an evaluation guide that contains scoring guides and key answer questions. This stage consists of 3 steps, namely: (1) Determination of the form of the instrument, (2) Preparation of the grid, and (3) Design of the instrument.

Development Stage

This stage consists of determining the form of the instrument, preparing a grid that is adjusted to the learning objectives and designing the instrument. The form of the test instrument developed in this study is a multiple choice test. Researchers chose the multiple choice form because it can be done easily, quickly and objectively. The question grid is used to determine the scope of the questions and instructions for making questions. After the grid is compiled, 20 scholastic test questions and 20 chemistry learning outcomes evaluation questions are developed.

In designing the instrument, it is done by making questions, scoring guidelines, and validation sheets. Making multiple choice questions. Making multiple choice questions is arranged based on the question instrument grid. Scoring guidelines are carried out by giving an assessment to each item, the correct answer gets a score of "1" while the wrong answer gets a score of "0". Making a validation sheet containing a questionnaire consisting of 21 questionnaire items with assessment categories in making validation sheets is divided into 4 categories, namely; (a) Material Aspects; (b) Construction Aspect; (c) Language Aspect; and (d) Additional Rules.
**Findings of Development Stage**

After the initial ability test instrument based on generic science skills is formed, it is in the development stage (develop) This instrument will be validated by expert validators and practitioners. The validation of the initial ability test instrument based on generic science skills in this study was carried out by 2 chemistry lecturers at Medan State University and 4 chemistry teachers. The questions that have been validated and revised are then tried out. The class that is the subject of the research is class XI D which consists of 36 students. The researcher gave test instrument sheets and work answer sheets then gave 30 minutes each to work on the scholastic test questions and the chemistry learning outcomes evaluation test. Small class tests were conducted to determine validity, reliability, level of difficulty, discriminating power and distractors.

The percentage (%) validation per aspect is obtained from the total score given by the validator divided by the maximum score that can be obtained in each aspect, then multiplied by 100%. After validating the test instrument, the material validation results are obtained as follows:

| Table 1. Lecturer and Teacher Validation Results for Each Aspect |
|---------------------------------|---------------|---------------|---------------|---------------|
|                                 | Material       | Construction   | Language       | Additional Rules |
| Teacher 1                       | 85             | 80             | 90             | 80             |
| Teacher 2                       | 87.5           | 95             | 100            | 100            |
| Teacher 3                       | 97.5           | 100            | 100            | 100            |
| Teacher 4                       | 85             | 90             | 75             | 100            |
| Lecturer 1                      | 82.5           | 87.5           | 95             | 100            |
| Lecturer 2                      | 85             | 90             | 95             | 100            |
| Rate (%)                        | 87.08          | 90.41          | 92.5           | 96.67          |
| Percent Interpretation Validation Criteria | Very high | Very high | Very high | Very high |

Based on Table 4.2 above which is the result of material validation, it can be seen that the average percentage results obtained in the material aspect are 87.08%, construction aspects are 90.41%, language aspects are 92.5% and additional rules are 96.67%. This shows that the initial ability test instrument is based on generic science skills on thermochemical material "worthy" used.

The results of the test instrument data analysis are as follows:

**Test Instrument Validity**

Testing the validity of the test is carried out with the aim of showing the level of validity or validity of the instrument to be used in research. In this study, the validity of the test instrument in terms of content quality was carried out by an expert validator, while to measure the validity of each item the test was administered to students. The number of students (N) who were the subject of analysis was 36 people. The data is obtained by using the product moment correlation formula. Meanwhile, from the data obtained r_y (validity) of each question item. The assessment criteria are if r_y > r_table then the question item is valid. Of the 20 questions on the scholastic potential test that were tested, there were 19 questions that were valid and considered appropriate, while 1 other question was invalid. Of the 20 chemical test questions tested, 17 questions were valid and considered appropriate, while the other 3 questions were invalid.

**Test Instrument Reliability**

The test reliability test is used to see how far the measuring device is reliable and trustworthy, so that the instrument can be accounted for in disclosing research data. A test is said to be reliable if the scores obtained by participants are relatively the same even though repeated measurements are taken. The reliability test in this study used the Kuder Richardson-21 (KR-21). From the product moment value tables, it is known that the value of r_table for N=36 and at =0.05 that is r_table = 0.329 while r_count = 0.769 (scholastic test) and r_count = 0.727 (chemistry learning outcomes evaluation test). By comparing r_count and r_table can be determined the reliability of test items, if r_count > r_table concluded reliable data. In the scholastic test instrument obtained r_count > r_table (0.769 > 0.329) it can be concluded that the test instrument is reliable. In the chemistry learning outcomes evaluation test instrument, it is obtained r_count > r_table (0.727 > 0.329) it can be concluded that the test instrument is reliable. So the questions can be used.

**Test Instrument Difficulty Level**

A good test item is a test item that is neither too easy nor too difficult (Silitonga, 2011). The results of the test for the difficulty level of the
Development of Initial Ability Test Instrument based on Science Generic Skills

A scholastic test instrument showed that out of 20 questions, there were 19 questions in the easy category, and 1 question in the medium category. The results of the difficulty level test of the chemistry learning outcomes evaluation test instrument showed that out of 20 questions there were 15 questions in the easy category, and 5 questions in the medium category.

Different Power of Test Instruments

Different Power of Test Instruments is the ability of a question to differentiate between students who are smart (high ability) and students who are less intelligent (low ability). The different power data of the questions can be seen in Figure 2 and Figure 3.

![Figure 2. Scholastic Question Distractors](image)

![Figure 3. Final Evaluation Question Distractor](image)

Different power aims to distinguish between high-ability students and low-ability students. Based on the calculation of the discriminating power of the scholastic test questions, it showed that there were 18 items that had good discriminatory power, and 2 questions had bad discriminatory power, while in the chemistry learning outcomes evaluation test instrument, 18 questions were obtained with good discriminatory power, and 2 questions with bad discriminatory power.

Test Instrument Distractors

Each multiple choice test has one question and multiple answer choices. Among the available answer choices, only one is correct. Besides the correct answer, there is a wrong answer. That wrong answer is known as a distractor (trickster). A distractor can be said to work well if it is selected by at least 5% of test subjects. The question distractor data can be seen in Figure 4 and Figure 5.

![Figure 4. Scholastic Test Different Power](image)

![Figure 5. Final Evaluation Question Distractor](image)

The questions developed are multiple choice questions with five answer choices, based on the results of the instrument test on the scholastic test item distractor it is known that out of 80 distractors there are 48 distractors that function properly, and 32 distractors that do not function properly and need to be revised. Based on the test results of the instrument distractor test item on the chemistry learning outcomes evaluation test, it is known that out of 80 distractors there are 56 distractors that function properly, and 24 distractors that do not function properly and need to be revised.

Of the 20 scholastic test items validated by students, it turned out that there were 2 items that could not be used, namely question number 14 (invalid with a validity level of 0.112 and poor discriminating power with a value of 0.00) while question number 19 had poor discriminating power. ie with different power levels of 0.1. There are 18 other questions that can be used with a validity level of 0.338–0.722, with an easy difficulty level (0.70–1.00) of 17 questions and 1 question with moderate criteria (0.30–0.69). Of the 18 questions that were declared valid, there were 8 questions that had good discriminating power (0.39–0.70), 9 questions with fairly good criteria (0.20–0.39). The reliability test result was 0.769.
in the high category (0.60 < r < 0.80 high reliability). Based on the test results, 18 questions were obtained that were suitable for use.

Out of the 20 scholastic test items validated by students, it turned out that there were 3 invalid questions, namely questions 16, 18 and 19 with sequential validity levels (0.70-1.00) of 11 questions and 5 questions with moderate criteria (0.30-0.69). Of the 18 questions that were declared valid, there were 6 questions that had good differentiability (0.39-0.70), 11 questions with fairly good criteria (0.20 - 0.39). The reliability test results were 0.727 in the high category (0.60 < r < 0.80 high reliability). So there are 17 questions that can be used.

Conclusion

This study aims to develop an initial ability test instrument based on science generic skills in the form of a scholastic test and develop a final chemistry evaluation test instrument in the topic of thermochemistry as well as find out the validity of the content and construction validity of the instrument.

Based on the results of the content validity test by experts, namely 2 lecturers at Medan State University and 4 chemistry teachers, the average percentage obtained in the material aspect was 87.08%, the construction aspect was 90.41%, the language aspect was 92.5% and the additional rule was 96.67%. This shows that the initial ability test instrument based on generic science skills on thermochemical material is feasible to use.

Based on the construction validation test data, it was found that of the 20 questions on the scholastic potential test tested, there were 19 questions that were valid and considered appropriate, while 1 other item was invalid, while of the 20 questions on the chemistry test tested, there were 17 questions that were valid and considered feasible, leaving 3 questions other invalid.

Acknowledgment

The researcher would like to thank all parties who helped complete this research, lecturers at Medan State University, SMAN 2 Percut Sei Tuan. Hopefully the results of this research can make a positive contribution in the field of education.

References


Irawan, A., & Kencanawaty, G. (2017). Peranan Kemampuan Verbal Dan Kemampuan...
Development of Initial Ability Test Instrument based on Science Generic Skills

Ambarita et al.


