

# Analysis of Validation of High-Thinking Valuation Instruments on Reproductive System Materials for Class XI High School

Lina Yuliyanti<sup>1,2,\*</sup> Yuni Ahda<sup>1</sup>

<sup>1</sup> Student of Master Program Biology Education Math and Natural Science Faculty, Universitas Negeri Padang, Padang, Indonesia

<sup>2</sup> Biology Department, Math and Natural Science Faculty, Universitas Negeri Padang, Padang, Indonesia

\*Corresponding author. Email: [Ylinayulianti@gmail.com](mailto:Ylinayulianti@gmail.com)

## ABSTRACT

Based on the analysis that has been done on daily test questions, it is found that the level of questions given to students is at the level of C1 to C3. Levels C1 to C3 are still at the LOST level, so it is not in accordance with the demands of the curriculum to analyze. Interviews were also conducted with the teacher to obtain additional information and it was discovered that the teacher had done a KD analysis before evaluating learning, but the teacher was still having difficulty making High Thinking Valuation Instruments. In response to this problem, than developed a High Thinking Valuation Instruments. The purpose of this research is to produce a High Thinking Valuation Instruments on a reproductive system material that is valid, practical, and of good quality items. This type of research is the development using the Plomp model. The Plomp model consists of three stages of development, are: preliminary research, the prototype phase and the assessment phase. Data collection instruments used were observation sheets, validity sheets, practicality sheets by teachers and students, observation sheets used by observers to assess affective and High Thinking Valuation Instruments to measure students' higher-order thinking skills as well as knowing learning competencies in the realm knowledge. The results of the validation of the High Thinking Valuation Instruments were judged by four experts. The aspects assessed consist of the aspects of the appropriateness of content, linguistics, and presentation. The validation results showed an average of 78.31% with valid criteria. The results of the validity assessment show that the High Thinking Valuation Instruments has met the valid criteria, so it can be used for the final test as an assessment of learning outcomes.

**Keywords:** *Validation, assessment instruments.*

## 1. INTRODUCTION

In Indonesia the ability to think at a high level is still very low. The results of a survey conducted by the Organization for Economic Co-operation and Development (OECD) using the 2015 Inter-National Program for Student Assessment (PISA) test, education in Indonesia ranks 69 out of 76 countries that take the PISA test. Besides PISA there are also Trends in International Mathematics and Science Study (TIMSS). The

results of the TIMMS study followed by Indonesia in 2011 scored 406, this value did not reach the international average score of 500.

Based on percentages, students in Indonesia on average have the right answer on the problem of application and reasoning lower than the matter of understanding. According to Mullis (2012), aspects of understanding, applying, and reasoning in the realm of cognitive abilities such as those applied at TIMSS can be used to show the profile of students' thinking abilities. The low level of high thinking can occur because in the

learning process students are not trained enough to hone the ability to think at a higher level.

Trials were conducted on March 2, 2018 to see the level of thinking ability of high-level students in SMAN 3 Padang, SMAN 4 Padang and SMAN 13 Padang against 96 students. The trial uses 10 items of multiple choice questions in the form of high-level thinking skills on material that students have learned, namely the material of the coordination system.

A trial of high-level thinking questions on the material regulatory system, obtained an average value of students for SMAN 3 Padang is 50.3, while SMAN 4 Padang 43.4 and SMAN 13 Padang 40.6. Based on these data, the ability to think of high-level students is still categorized as low. This happens because students are accustomed to low-level thinking problems, so when faced with high-level thinking questions students find it difficult to answer.

This can be seen from the results of the analysis of daily test questions in three high schools namely SMAN 3 Padang, SMAN 4 Padang and SMAN 13 Padang. Researchers analyzed the daily test questions on Reproductive System material in the even semester of the 2016/2017 school year.

Obtained the results of the analysis of UH questions on the reproductive system material that has been done at SMAN 3 Padang the level of questions is at the level of C1 to C3. Likewise with SMAN 4 and SMAN 13 Padang. When viewed from the demands of the basic competence of the Reproductive System material, the teacher should use the assessment for evaluation starting from the C4 level because the KD demands are to analyze and apply. The results of an interview with one of the teachers who teach biology at SMAN 3 Padang, Mrs. Dra. Azhira, M.Pd, is known that the teacher has done KD analysis before evaluating learning. However, teachers have difficulty making high-level evaluation assessments.

Considering the teacher's low ability to make high-level thinking questions, it is necessary to develop higher-level thinking assessment instruments that can be used as guidelines by the teacher in conducting further assessments. The questions developed are expected to help students improve their ability to solve problems with C4-C5 cognitive levels.

## **2. LITERATURE REVIEW**

### **2.1. Validity**

A test instrument is said to be valid if the test can measure what you want to be measured. According to Purwanto (2011), validity is related to the ability to measure exactly what you want to measure. Basrowi and Siskandar (2012) say:

Validity is a process carried out by the compiler or instrument user to collect data empirically to support the conclusions

generated by the instrument score. In other words, validity is a measuring tool to measure the measurement objectives".

According to Jihad and Haris (2012), validity aims to determine the suitability of the questions with the teaching material with the objectives to be measured or with the lattices that we make. Validity is done by requesting consideration of experts (experts) in the field of evaluation or experts in the field being tested. According to Sudjiono (2009), a test is said to have validity if the test is correct, correct, and valid or valid has been able to reveal or measure what is to be measured through the test.

Broadly speaking there are two kinds of validity, namely logical validity and empirical validity.

#### 1) Logical validity

The logical validity of an assessment instrument refers to the condition of the instrument which is said to be valid based on the results of reasoning. Basrowi and Iskandar (2016) state that the logical validity for an evaluation instrument refers to the conditions for an instrument that meets valid requirements based on the results of reasoning. The valid condition is considered fulfilled because the instrument in question has been well designed following the existing theories and provisions.

There are two kinds of logical validity that can be achieved by an instrument, which is as follows.

##### a) Construct validity

According to Sudaryono (2012), a test of learning outcomes can be stated as a test that has construct validity if the test of learning outcomes is viewed in terms of arrangement, framework or fiction has been able to accurately reflect a construction in psychological theory.

##### b) Content validity

Content validity is performed to show that a test instrument can reveal the contents of a concept or variable to be measured. According to Sudjana (2009), content validity is related to the ability of assessment tools to measure the content that should be. According to Sudaryono (2012), the content validity of a learning achievement test is the validity obtained after analyzing, tracing or testing the content contained in the test.

Qualitative analysis of the questions on the assessment instruments can be done by considering the content validity and construct validity. According to Permendikbud number 104 article 10 paragraph (1) of 2014, that the assessment instruments for knowledge competence contain at least material, construction, and language. Based on this, the

aspects analyzed qualitatively include material, construction, language / culture, and high-level thinking skills.

Logical validity is obtained from the validation results in the form of a questionnaire that has been filled out by the validator. Logical validity can be obtained directly without having to test it to students. This agrees with Arikunto (2012), that logical validity does not need to be tested for conditions, but is obtained immediately after the instrument has been prepared.

## 2) Empirical validity

Empirical validity is the same as criteria validity because validity is determined based on criteria. Putra (2013) argues, that empirical validity is validity that aims to measure the accuracy of an evaluation tool based on certain criteria.

Empirical validity is divided into 2 types, namely as follows.

### a) Predictive validity

Supardi (2015) states that an item of assessment instrument is said to have predictive validity or forecast validity if it has the ability to predict what will happen in the future.

### b) Validity of "existing now" (concurrent validity)

The validity of "existing now" (concurrent validity) is often known as empirical validity. According to Purwanto (2009), a test is said to have concurrent validity if the test results have a high correlation with the results of another measuring instrument against the same field at the same time.

Based on research conducted by researchers related to the results of a test that must have a high correlation with other measurement results on the same field at the same time, the empirical validity used in this study is the validity of the present (concurrent validity) which is more commonly known as empirical validity.

After completing the assessment instrument, an evaluation of the instrument is needed. This evaluation is intended to find out whether teaching materials have been good or if there are still things that need to be improved. The evaluation component includes the appropriateness of content, linguistics, and presentation. The component of content eligibility includes, among others:

- 1) Compliance with SK, KD
- 2) Compliance with child development
- 3) Conformity with the needs of teaching materials
- 4) The truth of the substance of learning material
- 5) Benefits for adding insight

- 6) Conformity with moral values, and social values

The Language Components include:

- 1) Readability
- 2) Clarity of information
- 3) Conformity with good and correct Indonesian language rules
- 4) Effective and efficient use of language (clear and concise)

The Presentation Components include:

- 1) Clarity of objectives (indicators) to be achieved
- 2) Order of presentation
- 3) providing stimulus
- 4) Completeness of information (depdiknas, 2008)

## 2.2. Higher-order thinking skills

According to Heong et al, (2011), higher-order thinking uses thinking widely to find new challenges. High-level thinking requires a person to apply new information or knowledge that he has acquired and manipulated information to reach possible answers in new situations.

Lewis and Smith (1993) define higher order thinking skills as thinking skills that occur when a person takes new information and information already stored in his memory, then connects that information and conveys it to achieve the desired goal or answer.

Stein (1996) in Tony Thomson (2008) defines high-level thinking is high-level thinking using complex, non-algorithmic thinking to complete a task, some cannot be predicted, using a different approach to an existing task and different from the example.

Senk, et al. (1997) in Thomson (2008) explained that high-level thinking is the ability to complete tasks where no algorithm has been taught, which requires justification or explanation and may have more than one possible solution.

According to Wardana in Rofiah (2013), high-level thinking skills are thought processes that involve mental activities in an effort to explore complex, reflective, and creative experiences carried out consciously to achieve goals, namely gaining knowledge that includes levels of analytical, synthesis, and evaluative thinking. According to Dewanto in Amalia (2013), the ability to think at a higher level is a capacity above the information provided, with a critical attitude to evaluate and

have metacognitive awareness and have problem-solving abilities.

Barnett & Francis (2012) states that higher-order thinking questions can encourage students to think deeply about subject matter. Higher-order thinking assessment instruments can provide stimulation because assessment in learning to develop high-level thinking of students. According to Wang and Wang (2011), there are three main components in HOTS, namely critical thinking skills, design thinking skills, and systems thinking skills. While Miri, et al. (2007), states that HOTS consists of three components, namely critical thinking skills, systematic thinking abilities, and creative thinking skills. According to King, et al. (2010) HOTS consists of two components namely critical and creative thinking skills.

The development of instruments for consideration HOTS is important in learning because the assessment of learning achievement changes as reforms throughout the world, especially in science education, promote a shift from traditional teaching to algorithmic, lower thinking abilities to, to higher-thinking skills (Barak and Dori, 2009 ). HOTS which aims to make students solve problems is something that is needed by students, so that when students graduate from the level of education students are ready to face greater challenges (Lailly, 2015).

### 3. RESEARCH METHODS

This type of research is the development using the Plomp model. The Plomp model consists of three stages of development, are : preliminary research, the prototype phase and the assessment phase. Data collection instruments used were observation sheets, validity sheets, practicality sheets by teachers and students, observation sheets used by observers to assess affective and and high-level thinking assessment instruments to measure students' higher-order thinking skills as well as knowing learning competencies in the realm knowledge.

The type of data that will be used in this study is primary data that is direct data obtained from the sample studied. This primary data obtained by students' learning competence in biology in the realm of knowledge and attitudes in the sample class.

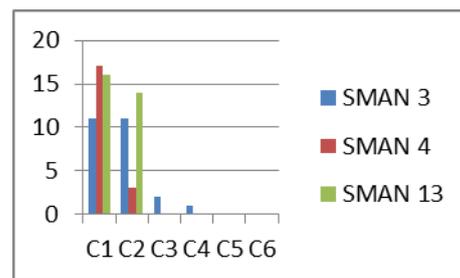
### 4. RESULT AND DISCUSSION

#### 4.1 Result

The results of the research carried out starting from the Preliminary Research Phase (Preliminary Research). At this stage a needs analysis, curriculum analysis and analysis of

students are aimed at obtaining a picture of the characteristics of the product being developed and can be used in the learning process. The activities carried out at the initial investment stage are Problem and Needs Analysis, student analysis, and curriculum analysis.

Problem analysis aims to find out the basic problems faced by teachers and students. The thing analyzed is the suitability of daily test questions with students' high-level thinking ability based on the 2001 Revised Bloom Taxonomy. The results of the question analysis are seen in Graph 1 for the Reproductive System material.



Graph 1: Results of analysis of daily repetition questions of the Reproduction system material

From Graph 1 above, it can be seen that the cognitive levels used in the three schools are still at the C1 and C2 levels, while C3 and C4 are only at one school, while the KD demand is the C4 level. The use of evaluation questions that are not in accordance with the KD demands, it is necessary to develop evaluation instruments that are in accordance with KD demands. In KD the reproductive system starts from the C4 level, and in the KD the immune system starts from the C3 level.

After the needs analysis is carried out, it is continued with the students' analysis which aims to find out the students' abilities. This characteristic can be seen from the value of the results of the evaluation of the ability to think at a higher level by giving tests of 10 items with the criteria of higher order thinking questions. The results of the evaluation of the ability to think at a high level of material Regulatory system that can be seen in Table 1.

**Table 1. Average evaluations of higher-order thinking skills**

School	Average value
SMAN 3	50,3
SMAN 4	43,4
SMAN 13	40,6

Curriculum analysis is also needed in making higher order thinking assessment instruments. Curriculum analysis is conducted by referring to Basic Competencies (KD) and indicators that are in accordance with the 2013 Curriculum. Curriculum analysis is used as a basis for designing higher-level thinking skills assessment instruments. The author develops this high-level thinking instrument in two KD namely 3.12 and 3.13.

The results of the analysis on KD 3.12 students are required to be able to Analyze the relationship between the structure of tissues making up the reproductive organs and their functions in the process of human reproduction through literature studies, observations, experiments, and simulations. The results of the analysis in KD 3.13 students are required to apply an understanding of the principle of human reproduction to cope with population growth through family planning (KB) programs and improving the quality of life of human resources through exclusive breastfeeding. Researchers do the development of pre-existing indicators by adding a number of suitable topics to be included in the indicators that the authors will use as guidelines in making instruments of high-level thinking assessment. The indicators that the authors add are about analysis of contraceptives. After the KD analysis the indicators and learning objectives are described. Indicators formulated are Analyzing the structure and function of the constituent tissues of male and female reproductive organs, comparing the stages of the process of spermatogenesis and oogenesis, analyzing the process of ovulation and menstruation as well as the factors that influence it, distinguishing the stages of fertilization, gestation and labor, analyzing the importance of ASI, analyzing the importance of KB to improve the quality of human resources.

Analyzing contraception, Analyzing Abnormalities / diseases that occur in the reproductive system. After formulating the indicators, the next step is to design an assessment of high-level thinking questions. The design stages are as follows.

- 1) Make a problem grid
- 2) Format selection
- 3) Designing the question cover
- 4) Design the description in the introduction.

- 5) Designing basic competencies, indicators, achievement of competencies and learning objectives
- 6) Question design guidance
- 7) Display high-level thinking assessment instruments
- 8) Design student answer sheets
- 9) Making an answer key

A high-level thinking assessment instrument that has been designed, then carried out by self-evaluation using a checklist to check for errors that might be found on the instrument. This checklist aims to produce instruments that are better than the initial design when given to the validator.

The next stage of the instrument will be validated by the lecturer as an expert validator. In addition to providing validator ratings, it also provides input for improvements to the high-level assessment instruments that have been designed. The high-level thinking assessment instrument that has been revised according to the validator's suggestion is then validated using a validation sheet. The aspects assessed in the validation of the high-level thinking assessment instrument by the validator include aspects, content eligibility, language and presentation aspects. The average score of validation results is 78.54% with a very valid category. The scores obtained indicate that the developed high-level thinking instrument was valid for use in learning. The results of the validation of the three aspects are listed in Table 2.

**Table 2. Results of Validation Analysis of Higher Level Thinking Assessment Instruments**

No.	Rated aspect	Validator				Validation Value (%)	Criteria
		1	2	3	4		
1.	Content Feasibility Aspect						
	The questions are arranged according to the core competencies found in the 2013 curriculum.	4	3	-	-	87,5	Very Valid
	The questions are arranged according to the basic competencies found in the 2013 curriculum.	4	3	-	-	87,5	Very Valid
	Questions are arranged in accordance with the development of students.	3	3	-	-	75	Valid
	Questions are arranged according to the needs of students.	3	3	-	-	75	Valid
	Questions are	3	3	-	-	75	Valid

	arranged to be able to broaden students' insights.						
	Problems measuring cognitive reasoning (analyzing, evaluating and creating)	3	3	-	-	75	Valid
	<b>Average</b>					<b>79,16</b>	<b>Valid</b>
2.	<b>The Aspect of Badness</b>						
	The description of the questions contained in the problem is easily understood.	3	3	4	4	87,5	Very Valid
	The language used is in accordance with the correct Indonesian language rules.	4	3	3	3	81,25	Very Valid
	Questions do not use sentences that lead to ambiguous.	3	3	3	3	75	Valid
	The language used in	3	3	3	3	75	Valid

	accordance with the level of understanding of students.						
	<b>Average</b>					<b>79,68</b>	<b>Valid</b>
3.	<b>Presentation Aspects</b>						
	The question cover has a clear identity	4	3	4	3	87,5	Very Valid
	Figures, graphs, data and tables are displayed accordingly	3	3	3	3	75	Valid
	The order of presentations in the questions is arranged systematically.	3	3	3	3	75	Valid
	The order of questions from low to difficult	3	3	3	3	75	Valid
	The suitability of the answer with the problem described.	3	3	3	3	75	Valid
	Accuracy of answers to the	3	3	3	3	75	Valid

	problems described						
	The length of the answer formulation is relatively the same	3	3	3	3	75	Valid
	<b>Average</b>					<b>76,79</b>	<b>Valid</b>
	<b>Total</b>					<b>235,63</b>	<b>valid</b>
	<b>Average</b>					<b>78,31</b>	

## 4.2. Discussion

Validation includes three aspects, namely the appropriateness of content, linguistic aspects, and presentation aspects. This validation is the stage to evaluate the instrument that has been made whether it is valid or needs revision. In addition to validating the validator also provides suggestions, the suggestions given form the basis of researchers to make improvements to the instrument to be more perfect. According to the Depdiknas (2008) The evaluation component includes the appropriateness of content, linguistics, offerings, and graphics. However, researchers only use three aspects according to their needs. Based on the validation by experts, it is known that the high-level thinking assessment instrument developed has met the valid criteria.

The validation assessment on the content eligibility aspect scored 79.16% with valid criteria. According to Geogory (2000 in Pamri 2011) content validity shows the extent to which questions, assignments or items in a test or instrument are able to represent the overall and proportional behavior of the sample subject to the test. These valid content requirements mean that the questions that have been developed already summarize the important components in the HOTS problem.

The results of the assessment on linguistic specs scored 79.68% with valid criteria. Language requirements are also very important in validator assessment. This is supported by Pamri (2011) which states that good

questions are questions that every item has to use language that is in accordance with good and correct language methods.

The results of the assessment on the aspect of the presentation obtained a value of 76.79% with a valid category. Rating of these three aspects received an average score of 78, 31% with a valid category. The score obtained shows that the developed high-level

## 5. CONCLUSIONS

The high-level thinking assessment instrument that has been developed is declared valid by a validator with valid criteria. Valid categories given by experts are based on the aspects of content, language and presentation. Based on the assessment, high level thinking assessment instruments can be used as innovations in the final evaluation of learning so as to be able to measure students' high thinking abilities.

## REFERENCES

- [1] Arikunto, s. 2013. *Prosedur penelitian: suatu pendekatan praktik*. Jakarta: rineka cipta
- [2] Barnett, J. E dan Francis, A.L. 2012. Using higher order thinking questionsto foster critical thinking: a classroom study. *International Journal of Experimental Educational Psychology*.
- [3] Basrowi dan iskandar. 2012. *Evaluasi belajar berbasis kerja*. Bandung : kerya putra darwati
- [4] Depdiknas. 2008. *Panduan Pengembangan Bahan Ajar*. Jakarta: Direktorat Jenderal Manajemen Pendidikan Dasar dan Menengah.
- [5] Jihad, asep dan Abdul Haris. 2012. *Evaluasi pembelajaran*. Yogyakarta: Multi Presindo
- [6] Heong, Y., dkk. 2011. The Level of Marzano Higher Order Thinking Skills Among Technical Education Students . *International Journal of Social and humanity*, Vol. 1,No. 2, July 2011.
- [7] King, JF, dkk. 2010. *Higher Order thinking skills, Definition, Teaching Strategis, Assesment*. A Publication of theeducational Services Program. (online). Diakses 22 Februari 2018.
- [8] Lailly N R and Asih W W. 2015. Analisis soal tipe Higher Order Thinking Skill (HOTS) dalam soal UN Kimia SMA rayon B Tahun 2012/2013 *Jurnal Kaunia*
- [9] Lewis, A & Smith, D. (1993). *Defining Higher Order Thinking*. Theory Into Practice, *International Journal*. Vol 32(3), 131-137
- [10] Mullis, I V.S, et al. 2012.Improving the Ability of Creative Thinking Mathematically and Self-Confidence Student through Application Model Eliciting Activities (MEAs) Review from Student Gender. *American Journal of Educational Research*. 2018, Vol. 6 No. 4, 319-32
- [11] Purwanto. 2011. *Evaluasi hasil belajar*. Yogyakarta. Pustaka belajar.
- [12] Sudaryono. 2012. *Dasar-dasar evaluasi pembelajaran*. Yogyakarta: Graha Ilmu
- [13] Sudjana, 2009. *Penilaian hasil proses belajar mengajar*. Bandung: PT Remaja Rosdakarya
- [14] Permendikbud. Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia No. 67 Tahun 2013 tentang Kerangka Dasar dan Struktur Kurikulum Sekolah Menengah Pertama/Madrasah Tsanawiyah.
- [15] Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia No. 68 Tahun 2013 tentang Kerangka Dasar dan Struktur Kurikulum Sekolah Menengah Pertama/Madrasah Tsanawiyah.
- [16] Plomp, T dan Nieveen. 2013. *Educational Design Research dalam Educational Design Research: An Introduction*. Enschede, University Of Twente.
- [17] Wang, S, dan Hai W. (2011). "Teaching Higher Order Thinking in the Introductory MIS Course: A Model-Direct Approach." *Journal Education for Business*, volume 86: 208-212.