THE DEVELOPMENT OF ASSESSMENT INSTRUMENT USING TESLET MODELS IN STEREOCHEMISTRY MATERIALS

Annisaul Khasanah Wulandari[1], Jimmi Copriady[2], Rasmiwetti[3]

(annisaulkhasanahwulandari@gmail.com)

Postgraduate Department of Chemistry Education, Faculty of Teachers Training and Education, Universitas Riau, Pekanbaru, Indonesia

Abstract: Stereochemistry is the study of molecules in three-dimensional space regarding how atoms in a molecule are arranged in one room relative to the other. Stereochemistry is a difficult course. Therefore, it is necessary to diagnose the difficulty of learning using the testlet model. This study is aimed at developing an appropriate testlet model for the assessment instrument. This research uses the development model of Borg and Gall's assessment instrument with ten stages, but this research is limited to the fifth stage only, namely 1) Research and Information Collection, 2) Planning, 3) Develop Preliminary Form of Product, 4) Preliminary Field Testing, and 5) Preliminary Product Revision. The data collection techniques were in the form of validation questionnaires from expert teams and students' response questionnaires. The data analysis technique used qualitative and quantitative data. The qualitative data were from the results of a review of validation by a team of experts who are Postgraduate Lecturers of Chemistry Education at Universitas Riau and lecturers of the Chemistry Education Study Program at the Faculty of Tarbiyah and Teaching at UIN Suska Riau. The quantitative data were obtained from the calculation of the mean scores of the validation questionnaire by two lecturers of the Postgraduate of Chemistry Education at Universitas Riau and one lecturer in the Chemistry Education Study Program at the Tarbiyah Faculty and Teacher Training at UIN Suska Riau as the experts. The results of the study show that the mean score of validation from the expert team is 3.97 on the aspect of material with valid category, 3.97 on the aspect of construction with valid category, 3.88 on the aspect of language with valid category; and the results of the students' response questionnaires are 51.58 on the time readability questionnaire and 8.08 on the aspect of time adequacy.

Keywords: *development, assessment instrument, testlet model, stereochemistry*

INTRODUCTION

Education plays an important role in educating the nation, developing and improving the quality of human resources (HR). Universities are part of the formal education system that have clear rules as a reference for the learning process.
Chemistry is a science that is included in the group of Natural Sciences (IPA). Thus, chemistry has the same characteristics as the other natural sciences which are related to how to find out about natural phenomena in a systematical way so that not only is mastery of a collection of knowledge in the form of facts, concepts, or principles required but also a process of discovery.

Chemistry has various fields of science. One of the fields of chemistry is organic chemistry. Organic chemistry is the branching of scientific studies of chemistry regarding the structure, properties, composition, reactions, and synthesis of organic compounds. Stereochemistry is the study of molecules in three-dimensional space — that is, how atoms in a molecule are arranged in one room relative to the other (Ralp, J. Fessenden and Joan S. Fessenden, 1982). In universities, the course of organic chemistry is gi en through Stereochemistry materials. Stereochemistry material is considered difficult. Stereochemistry material requires the determination of students to read and understand the concepts and solve the questions.

Based on the interviews with lecturers who teach stereochemistry subjects, it is known that the tests used in the course are in the form of descriptive tests. Moreover, some students have difficulty learning in stereochemistry materials. The lecturers have never used assessment instruments (diagnostic tests) on organic chemistry, especially stereochemistry.

Based on interviews through questionnaires on 60 students consisting of 20 students from the Chemistry Education Study Program at the Faculty of Tarbiyah and Teaching at UIN Suska Riau and 40 students from the Chemistry Education at the Faculty of Teachers Training and Education at Universitas Riau, 100% stated that they liked organic chemistry but 70% stated organic chemistry was difficult and 75% stated that stereochemistry was difficult, with details: 8.33% did not understand alkane geometry, 38.33% did not understand cyclic geometric isomers, 65% did not understand conformational open-chain compounds, 71.66% did not understand conformation cyclohexane and 46.66% did not understand chirality. Thus, it is necessary to analyze the assessment instrument using a testlet model on stereochemistry materials.

Assessment is a systematic and continuous process or activity in gathering information about the process and outcomes of students learning in order to make decisions based on certain criteria and considerations (Arifin, Zaenal 2009). Assessments can be in the form of tests and in the form of non-tests.

A test is one form of instruments used to make measurements. Measuring instruments in the form of tests generally provide information about the cognitive characteristics of the test taker. Judging from its objectives, there are four types of tests that are widely used in educational institutions, namely: (1) placement tests, i.e. tests carried out at the beginning of the lesson and used to determine the students' level of ability, (2) diagnostic tests, i.e. tests useful to find out the learning difficulties faced by students, including misunderstanding of concepts, (3) formative tests, i.e. tests that aim to obtain input on the level of success of the learning process, (4) summative tests, i.e. tests given at the end of a lesson or end of semester to determine student learning success (Suwarto, 2013). The assessment instrument used in this study is in the form of a diagnostic test.

Diagnosis can be interpreted as an attempt to find out what weakness or disease someone is experiencing through careful testing and study of symptoms. Diagnosis can also be interpreted as something that finds characteristics or errors (Thorndike and Hagen in Syamsudin. A, 2002).

The purposes of diagnostic tests are: 1) to determine the strengths and weaknesses of students; 2) to determine the teaching that needs to be done in the future; 3) to be the tools and instruments used to identify learning difficulties; 4) to provide tools to find the causes of deficiencies.

Indah Wahyuni, et al. (2015) explain that testlet tests combine the advantages of multiple choice questions and problem essay questions. Testlet tests consist of main questions and supporting questions. The main questions are in the form of essay questions and supporting questions in the form of objective questions.

The testlet model aims to diagnose the strengths and weaknesses of students and to help determine the level of the students. The testlet model also aims to look at the individual profile of students.

RESEARCH METHODOLOGY

This study uses the Research and Development R & D method using the Borg and Gall development model with 10 steps, namely: 1) Research and Information Collection, 2) Planning 3) Develop Preliminary Form of Product 4) Preliminary Field Testing, 5) Preliminary Product Revision 6) Main Field Testing 7) Main Product Revision 8) Operational field testing - counted 9) Operational Field Testing Revision 10) Final product (dissemination and implementation). Figure 1 shows the steps of R & D Research and Development using the Borg and Gall development model.

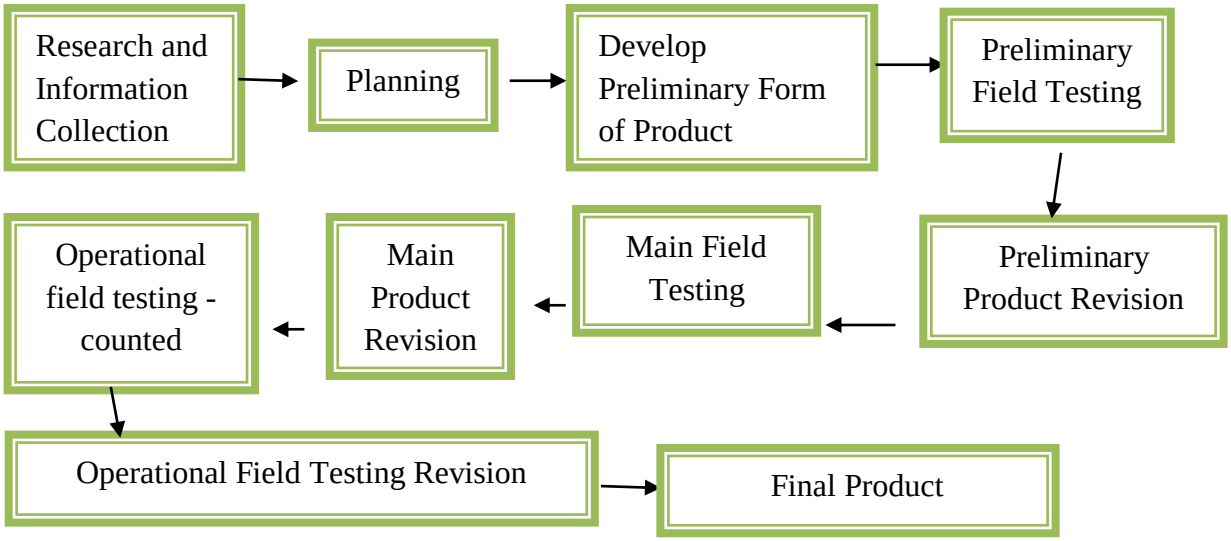
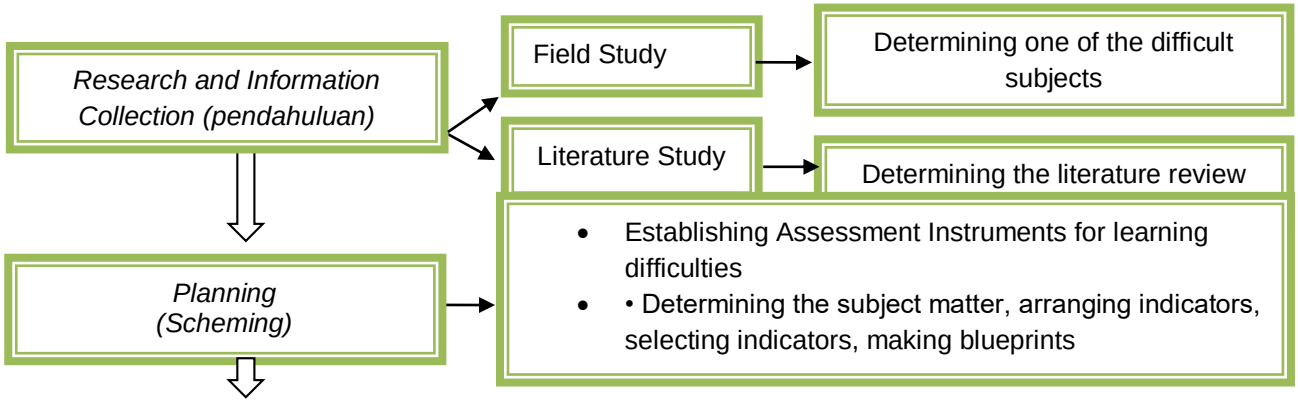


Figure 1: The steps of R & D Research and Development using the Borg and Gall (1989) development model.

However, this research is limited to the fifth stage only because it is adjusted to the time and needs of the study. Figure 2 shows the Procedure Scheme of Research that has been carried out until stage 5 based on the research scheme for the Brog and Gall development model.



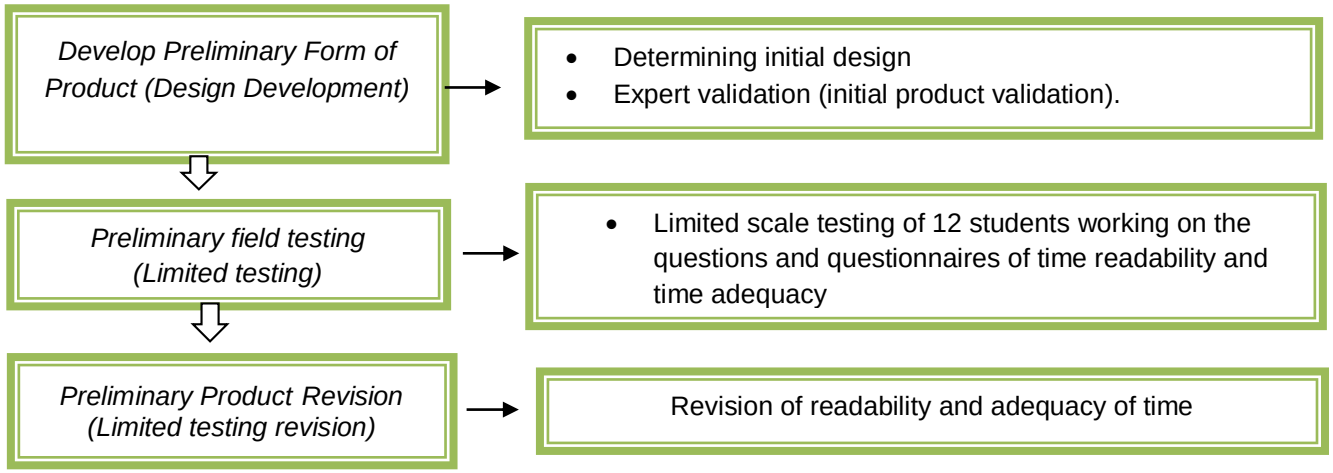


Figure 2: Research Procedure Scheme that has been carried out until stage 5 based on the research scheme using Brog and Gall development.

The data collection techniques gathered (1) qualitative data (data obtained through comments and suggestions from the three validators on the products developed) and (2) quantitative data (data obtained from the validation results in the form of calculating the mean score of item validity by two Postgraduate lecturers of Chemistry Education Study Program at Universitas Riau, and one lecturer of Chemistry Education Study Program at the Faculty of Tarbiyah and Teachers Training at UIN Suska Riau as the validators).

Arikunto states that in order to find out the final grade rating for each item in the research questionnaire, the number of scores obtained is divided by the number of respondents (validators) who answer the assessment questionnaire. Based on that opinion, the formula for calculating the mean score is as follows:

$$\bar{X} = \frac{\sum x}{n}$$

Where:

$\bar{X}$ = mean score
$\sum x$ = the sum of response scores
n = the number of validators

The criteria of the Likert scale validation used can be seen in table 1 below:

Table 1 Likert Scale for Material, Construction, and Language Validations

Mean Score	Validation Criteria
3,28 – 4,00	Valid
2,52 – 3,27	Quite valid
1,76 – 2,51	Less valid
1,00 – 1,75	Not valid

(Arikunto, 2016)

FINDINGS AND DISCUSSIONS

The results obtained from the research carried out based on the steps using Brog and Gall development are as follows:

1. Research and Information Collection (preliminary)

The stage of Research and Information Collection (preliminary) was using field study and literature study. The stage in the field study can be carried out through interviews.

At the stage of field study with lecturers who teach stereochemistry materials, some information is obtained as follows: (1) The type of test used for evaluation is a test in the form of an essay test; (2) Some students have difficulty in learning organic chemistry courses, especially stereochemistry materials which are reinforced by questionnaires; (3) The lecturers have never diagnosed students' learning difficulties on organic chemistry courses, especially stereochemical materials.

The results obtained at the preliminary stage by carrying out theoretical study activities are:

a. The results of the theoretical study with books obtain theories of development (definitions of research and development [R & D], defenitions of research and development [R & D] according to Borg and Gall), theories of assessment instruments (diagnostic tests), testlet models, and stereochemistry.
b. The results of the material study find out parts of the material used.
c. The results of the indicator study find out that the question indicator studies are in accordance with the materials and indicators in the organic chemistry syllabus about stereochemistry.

d. Researchers can construct assessment instrument questions on strerochemistry material using the teslet model.

2. Planning (Scheming)

The planning stage was done by determining the product and instrument, described as follows:

a. The determination of the products was done through discussions with supervisors and colleagues about the testlet questions being developed.
b. The Stage of Determining Instrument includes the design and preparation of supporting instruments for research and development of the assessment instruments using the testlet model. The results obtained from the design and instrument determination include:
    1) The material listed in the question blueprints
    2) Standard competency listed in the question blueprints
    3) Indicators and question indicators listed in the question blueprints
    4) Questions and key answers listed in the question blueprints
    5) Scores of teslet model listed in the question blueprints
    6) Question blueprints

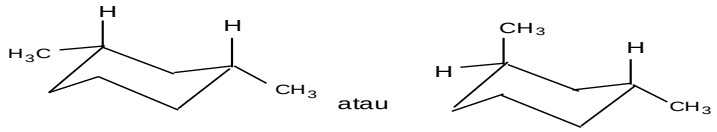
3. Develop Preliminary Form of Product (Design Development)

The design development stage was done through these activities:

a. Determining the initial design of the questions that are consulted with the consultants of peers and lecturers.
b. The development of the instrument assessment design using testlet model is validated by 3 validators; 2 persons from the Postgraduate Program of Chemistry Education at Universitas Riau and 1 person from Chemistry Education Study Program at the Faculty of Tarbiyah and Teachers Training at UIN Suska Riau as experts (validators).
c. The design of the questions has resulted in the assessment instrument which has been composed of 40 multilevel multiple choice questions (testlet) with 20 main questions that have 5 answer options each, and the material used is Stereochemistry. Figure 3 shows an example of an assessment instrument using a testlet model on stereochemistry material.

Informasi berikut ini digunakan untuk mengerjakan soal nomor 1 dan 2

Sandra membuat suatu pasangan konformasi 1,3-dimetil-sikloheksena yang terdisubtitusi seperti dibawah ini:



1. Isomer mana yang lebih stabil dari suatu pasangan konformasi sikloheksana yang terdisubtitusi adalah...
   a. Trans a,a lebih stabil dari trans e,a
   b. Cis, a,a lebih stabil dari trans e,a
   c. Trans e,e lebih stabil dari cis a,a
   d. Cis e,e lebih stabil dari trans a,e
   e. Trans, a,a lebih stabil daripada cis a,e

2. Pernyataan yang benar tentang stereokimia berdasarkan jawaban soal nomor 1 adalah...
   a. Karena pada trans a,a memiliki tolakan yang lebih besar maka energinya lebih tinggi dari pada cis e,a karena salah satu subtituen cis (ekuatorial,e) memliki tolakan yang lebih sedikit
   b. Karena pada cis e,e kedua subtituen ekuatorial (e,e) memiliki tolakan lebih sedikit maka mempunyai energi lebih rendah dari pada trans a,e karena salah satu subtituen trans (aksial,a) memiliki tolakan yang lebih besar
   c. Karena pada cis a,a kedua subtituen aksial (a,a) memiliki tolakan lebih besar maka mempunyai energi lebih tinggi dari pada trans a,e karena salah satu subtituen trans (ekuatorial,e) memliki tolakan yang lebih sedikit
   d. Karena pada trans e,e kedua subtituen ekuatorial (e,e) memiliki tolakan lebih sedikit maka mempunyai energi lebih rendah dari pada trans a,a karena kedua subtituen aksial (a,a) memiliki tolakan lebih besar
   e. Karena pada trans a,a memiliki tolakan yang lebih besar maka energinya lebih tinggi dari pada trans a,e karena salah satu subtituen trans (ekuatorial,e) memliki tolakan yang lebih sedikit

1) The data used in this research are qualitative and quantitative data.

Qualitative data were gained through comments and suggestions from the reviews by the three validators based on aspects of materials, constructions, and language. Table 2 shows the results of comments and suggestions from the reviews of the three validators.

Table 2: The results of comments and suggestions from the reviews of the three validators.

	Validation 1	**Validation 2**	**Validation 3**
Validator 1	• Some sentences should be revised in order to make it easy to understand, well-written and interesting • Some sentences in the test are still not in accordance with EYD (Enhanced Spelling)	The instrument is suitable for use	None

	Validation 1	Validation 2	Validation 3
	• Some typing errors in the questions should be fixed • The questions are broad and in accordance with the development of college students • The instrument is complete (question grid, question items and scoring guidelines)		
Validator 2	• Question number 1: Regarding the name of the compound (Z)-1-amino-2-hydroxy-2-methoxyethane, why is there amine (NH3) in the options? Please correct it! • Question number 2: (1) Replace the carbon chain to carbon. The carbon chain is like (CH3-CH2-CH3) while the carbon is C. (2) Replace amino to amina because the name of the compound is (Z)-1-amino-2-hydroxy-2-methoxyethane. • Question number 7: regarding 1-methoxy-2-methyl-4-isopropylcyclohexane, add "trans" in the beginning of it so that it is not difficult to distinguish the structure. • Question number 8: delete “cis” and “trans”. • Change the form of the structure on the information (9 and 10). • On questions number 9, 11, and 13, choose one of the structures on the options. • Recheck and understand the structure information (17 and 18; 19 and 20; and 21 and 22). • Change the options number 30. • Change the options numbers 33 and 34.	• Please check and undertand. • When you make a question test, you should understand the materials. • Test trial. Mr. Jasril asked and I answered correctly.	The instrument is suitable for use.
Validator 3	• The wording on some questions are not in accordance with EYD (1, 2, 3, 11, 12, 13, 15, 21, 29, 30).	• Adjust the structure options a b c	The instrument is

Validation 1	Validation 2	Validation 3
• Write the name of the structure. • Fix sentence number 12. • Add numbering on the essay test. • Adjust the numbering for the essay test and the objective one. (21, 22, 37, and 39)	d e with the name of the compound in the question. For example, the kloro nomenclatur e turns out to be in the bromo structure.	suitable for use.

The quantitative data were obtained from the calculation of the mean score of the validation assessment sheets which include 3 aspects of criteria: 1) aspect of material, 2) aspect of construction, and 3) aspect of language.

1) Aspect of material

The validation results of the assessment instrument using a teslet model on stereochemistry material on the aspect of material can be seen on Table 3.

Table 3: Validation on the aspect of material

No	Criteria	Validator's Score			Mean Score	Category
		1	2	3		
1.	The representativeness of testlet models with stereochemistry concepts	3,975	4,000	3,950	3,975	Valid
2.	The coverage of problem indicators with indicators of stereochemistry learning	4,000	4,000	4,000	4,000	Valid
3.	The conformity between the assesment instrument *of* learning *difficulity* with stereochemistry testlet models	4,000	4,000	3.925	3.975	Valid
4.	Material in assessment instruments learning difficulties	3,975	4,000	3,825	3,930	Valid
Average					**3.97**	**Valid**

2) Aspect of contruction

The validation results of the assessment instrument using a teslet model on stereochemistry material on the aspect of construct can be seen in Table 4.

Table 4. Aspect of construct.

No.	Criteria	Validator's Score			Mean Score	Catagory
		1	2	3		
1.	The readability of the *Assessment* Instrument on stereochemistry learning difficulties	3,975	4,000	3,85	3.940	Valid
2.	Conformity of the *Assessment* instruments of learning difficulties with the answers to the instruments	4,000	3,925	3.95	3.975	Valid
3.	The completeness of the *Assessment* instrument (question grid, question items, answers and scoring guidelines)	4,000	4,000	4,000	4,000	Valid
4.	The conformity of the questions broadness with the development of college students	4,000	4,000	3.950	3.980	Valid
Average					**3.97**	**Valid**

3) Aspect of language

The validation results of the assessment instrument using a teslet model on stereochemistry material on the aspect of language can be seen in Table 5.

Table 5: Aspect of language

No.	Criteria	Validator's Score			Mean Score	Catagory
		1	2	3		
1.	The presentation of problems and writings in the *Assessment* Instrument	3,950	4,000	3,675	3,875	Valid
2.	The presentation of language/wording in the *Assessment* Instrument	3,925	3.900	3,875	3,9	Valid
Average					**3,88**	**Valid**

2) Prelimery field testing (Limited Testing)

The subjects of the limited test or limited-scale field test were 12 students of the Chemical Education Study Program at the Faculty of Tarbiyah and Teaching at UIN Suska Riau.

The time provided to answer the limited scale test is 100 minutes for 2 credits. The determination of the test time for 100 minutes was considering the average duration of answering each question which is 2.5 minutes. This time is estimated to be enough to answer the question.

3) Preliminary Product Revision (Initial product revision/Limited test)

The data on the limited scale test are data in the form of questionnaires of readability and time adequacy in answering the questions and the test answers. The questionnaires distributed to 12 students are summarized and calculated to get conclusions about whether the question has a readability aspect and a good aspect of time adequacy. The average score of the readability aspect is 51.58 with excellent qualifications and the average value of time adequacy aspect is 8.08 and categorized well qualified. Thus, the questions can be used without revision and the time used is 2 credits (100 minutes) and can proceed to a large-scale testing phase.

CONCLUSION AND RECOMMENDATIONS

The conclusions are as follows:

1. Assessment instrument on learning difficulties in stereochemistry material used a testlet model developed with 10 steps of development by Borg and Gall but in this study with only 5 steps of the development, including: (1) Research and information; (2) Planning; (3) Develop preliminary form of product; (4) Preliminary field testing; and (5) Preliminary Product revision.
2. The results of the calculation of the mean score of the validation from the expert team are 3.97 on the aspect of material (valid); 3.97 on the aspect of construction (valid); and 3.88 on the aspect of language (valid).
3. The result of the student responses is 51.58 on the time readability questionnaire and 8.08 on the adequacy of time questionnaire.

The recommendations are as follows:

1. The development of assessment instruments on students' learning difficulties must pay attention to the time allocation of the test.
2. A testlet-modelled instrument can be used as an alternative development of assessment instruments to overcome student learning difficulties, especially in stereochemistry material.

REFERENCES

Aloysius Rabata Nova TFT, Parno, Supriyono Koes H. (2016). Pengembangan instrumen assessment penguasaan konsep tes testlet pada materi suhu dan kalor. *Jurnal pendidikan: teori, penelitian, dan pengembangan volume: 1 nomor: 6 bulan juni tahun 2016.* Universitas Negeri Malang. Malang

Arifin, Zaenal. (2009). *Evaluasi Pembelajaran*. Bandung. Remaja Rosdakarya.

Arikunto, S. (2014). *Prosedur penelitian suatu pendekatan praktek*. Jakarta. Rineka Cipta.

Arikunto, S. (2016). *Dasar-dasar evaluasi pendidikan*. Jakarta. Bumi Aksara.

Arsyad, Azhar (2000). *Media Pengajaran*. Jakarta. PT Raja Grafindo Persada.

Borg, W.R. & Gall, M.D. (1989). Educational Research: An Introduction, Fifthy Edition. New York.Longman.

Dimyati & Mudjiono. (2009). *Belajar dan Pembelajaran*. Jakarta. PT Rineka Cipta.

Djamarah, S.B. (2002). *Psikologi belajar*. Jakarta. Rineka Cipta.

Levi Nahum, Havi Hofstain, Rachel Mamlock-Naaman & Zira Bar-Dov. Can Final Examinitions Amplify Student' Misconception in Chemistry. *Journal Chemistry Education: Research and Practice. Vol 5 No 3 Pp.301-325.* Israel.The Weizmenn Institue of Science, Deparment of Science Teaching.

Putra, Nusa. (2015). *Research & Development*. Jakarta. Rajawali Pers.

Ralp, J. Fessenden & Joan S. Fessenden. (1982). *Kimia organic edisi ketiga*. Jakarta. Erlangga.

Indah Tri Wahyuni, Sri Yamtinah, Budi Utami. (2015). Pengembangan Instrumen Pendeteksi Kesulitan Belajar Kimia SMA Kelas X Menggunakan Model Testlet. *Jurnal Pendidikan Kimia (JPK), Vol. 4 No. 4 Tahun 2015.* Surakarta. UNS.

Suwarto. (2013). *Pengembangan tes diagnostik dalam pembelajaran panduan praktis bagi pendidik dan calon pendidik*. Yogyakarta. Pustaka Pelajar

Sri Yamtinah, Haryono, Kus Sri Martini. (2014). Profil Individu Peserta Didik Pelengkap Tes Jenis Testlet Sebagai Alternatif Pendeteksi Kesulitan Belajar Kimia. *Jurnal Profesi Pendidik. Volume 1, Nomor 1, November 2014*. Surakarta.UNS.

Syamsudin, A. (2002). *Psikologi Pendidikan: Perangkat Sistem Pengajaran Modul*. Bandung. PT Ramaja Rosdakarya