

DEVELOPING HIGHER ORDER THINKING SKILL TEST INSTRUMENT FOR THE 10TH GRADE ENGLISH MATERIAL AT MA MADANI ALAUDDIN PAO-PAO PAPER TITLE

Mardiana¹

Nur Asik²

Eril Sandi³

St. Aminah Putri Abdullah⁴

^{1,2,3,4}Universitas Islam Negeri Alauddin Makassar

e-mail: mardiana@uin-alauddin.ac.id

ABSTRACT

The purpose of this study is to develop HOTS items in the "Both You and I" material, to determine the level of validity of the items, to determine the reliability of the items for class X in MA Madani Alauddin Pao-Pao. The type of this research is research and development (R&D). The development model that is used, refers to the Plump model which consists of five stages of development, namely Preliminary investigation, Design, Realization/Construction, Test, Evaluation and Revision (Test, Evaluation and Revision), and Implementation (Implementation). The subjects in this study are students at class X MIA 2 in MA Madani Alauddin Pao-Pao. The instruments are used in data collection are validation sheets and HOTS items. Based on the results of the study, it obtained that the level of validity of the HOTS items was 3.7 which is in the very valid category, the reliability of the HOTS items is 0.81 shows in very good category, the difficulty level of the HOTS items is 0.60 shows in the medium category. The implications of the developing HOTS items it can train and encourage students to do high-level reasoning so that they are not fixated on the pattern of answers that result from the memorization process without knowing the concept of knowledge, and it can make students think critically in working on the test. So that the answers can be produced based on in-depth reasoning from the questions for each item, namely 10 multiple choice questions and 5 essay questions.

Keywords: Higher Order Thinking Skill Test, Instrument Test, Bank Soal, Wordpress

1) INTRODUCTION

Nowadays, the curriculum that is implemented by the government is the 2013 curriculum (K13), where the principle of this curriculum is to prepare human resources to have the ability to live as individuals and citizens who are faithful, productive, creative, innovative, and effective to produce Indonesian people through the formation and strengthening of aspects of Attitude, Skills and Knowledge that are integrated. Also be able to contribute to the life of society, nation, state and world civilization (Widana: 2017).

Widana (2017), stated that the result of the research by Program for International Students Assessment (PISA), data in 2015 declare that Indonesia is in the bottom 4 of the 43 participated countries in the world. Indonesia has low rank in reading literacy, mathematic literacy, and scientific literacy. It is shown by answering the question of the test at the research. This concludes that the objective of the education in Indonesia is not relevant to the objective of 2013 Curriculum. The Government (2017) also mentioned that the student's ability is lack in understanding complex information, understanding theory and solving problem. So that the government has a new revision in Curriculum.

The curriculum has revised by Ministry of Education and Culture of Indonesia which known as 2013 Curriculum revised edition, the 2013 curriculum is revised with some completing. The completing such as the implementation on the content of the standard is minimizing the materials that are not relevant with the development of the material for student's need to thinking critically, the change of context standard, and the assessment standard, also the assessment standard in this revision is assessing the student's result of learning process that helping the student to increase the HOTS, because higher thinking can pull the student to think wider and more deeply about the learning (Widana, 2017).

The purpose of education will be successful achieved if the implementation is released that based on government's certainty. It is following the assessment system for the students based on the government's rules. The assessment research of education consists of the result of study by the teacher, the result of the study by the unit of educator and the result of study by the government. The result of the learner's assessment that it is seen by the teacher by using some technique such as test, observation, individual task or group, and the other form which appropriates with the characteristic of the competency and the stage of the learner's development (Salamah, 2018:274).

A test as an instrument of learning outcomes should be measured student skills at various levels, ranging from low levels of thinking to high levels of thinking. Therefore, it is necessary to pay attention to each proportion of the level of thinking ability in each question problem that will affect student learning patterns. the fact, in making instruments questions tend to be dominated by questions with low-level thinking skills so students prefer to learn by memorizing and working on the questions contained in the book rather than developing their ability at a higher level of thinking when trying to solve a problem.

Problems that occur in schools, questions tend to be more numerous testing aspects of memory that lack training in higher order thinking skills students, the ability to think of Indonesian children scientifically is still considered low seen one of the factors causing it, among others because students in Indonesia are less trained in solving problems that are measure HOTS, and the problem faced by the teacher is the ability of the teacher in developing HOTS assessment instruments are still lacking and not yet the availability of assessment instruments specifically designed to train HOTS, so that HOTS assessment instruments need to be developed. Development of thinking skills a high level of students will result in: students' proficiency in problem solving strategies become good, the level of confidence of students in mathematics increases, and students' learning achievement on non-routine problems which requires higher-order thinking skills to increase.

The students need to increase their thinking skills by giving the instrument of test based on the HOTS concept in 2013 curriculum. From the explanation previously, the researcher considers to conduct the research about design the instrument test based on the English material which is applied HOTS concept.

2) METHODS

Research Method

In this research, the researcher used Research and Development (R&D) method. R&D is defined as a research design that involves the classroom problems, studying recent theories of educational product development, developing the educational products, validating the product to experts, and field testing the product (Latif, 2012).

R&D is defined as a research design that involves the classroom problems, studying recent theories of educational product development, developing the educational products, validating the product to experts, and field testing the product (Latif, 2012).

It consists of a number of steps commonly referred to as the "R&D cycle", which involves researching the research findings related to the product to be developed, developing the product based on the findings, conducting field tests in the environment where it will eventually be used, and making revisions that can be corrected found during the field test phase. (Borg, Gal and Gal: 2003).

In addition, educational research and development (R&D) is also defined as a process aimed at developing and verifying educational products. (Borg, Gal and Gal: 2003). It consists of a number of steps commonly referred to as the "R&D cycle", which involves researching the research findings related to the product to be developed, developing the product based on the findings, conducting field tests in the environment where it will eventually be used, and making revisions that can be corrected Defects found during the field test phase. (Borg, Gal and Gal: 2003).

In research and development, the researcher used the Tjeerd Plomp model as one of the models. The Plomp model is shown to be more adaptable and versatile than the Four-D model since it includes development activities at each step that may be tailored to the peculiarities of the research. Preliminary research, for example, may be included in the preliminary investigation. This preliminary research can also take the form of preliminary research whose findings serve as the foundation for future development; in this case, the research design is a research and development design (Rochmad,2012).

Plomp (1997: 5) said, "*we classified educational design in brief as a strategy within which one works in a systematic way towards the solution of a make problem.*" Characteristics of design in the field of education as a method through which people work in a systematic manner to solve problems. Penelitian design focuses on the Plomp model, which is commonly used by students, including S1, S2, and S3 students, when conducting research. The best model for resolving the problem of Plomp's education (1997: 5).

This methodology has five stages: (1) preliminary investigation, (2) design, (3) realization / construction (realization / construction), (4) test, evaluation, and revision (test, evaluation, and revision), and (5) implementation (Implementation). The Plomp development model was chosen because it is more versatile, as each step involves bespoke development activities based on the research's characteristics.

Research Subject

This research will take place at MA Madani Alauddin Pao-pao. It will involve several research participants. They are selected based on Purposive sampling in order to conduct data collection Therefore, this research employs multiple data sources by inviting learners and teachers of MA Madani Alauddin Pao-pao and expert. To be more detail, the research subjects are elaborated as follows:

1. Students (Target Group): 20 students will be the subject or the people whom this development is addressed to.
2. Teachers (Audience Group): 2 teachers.
3. Experts: 2 experts be involved in this research due to make sure the appropriateness of the materials.

Research Instrument

The results of conducting this study requiring data. The researcher use some following tools to obtain the data as follows:

1. File

During the research process the researchers collected qualitative documents. Specifically, the literature in this study is to train teachers to train students' higher-order thinking ability. In the assessment of high-level thinking skills, it is necessary to understand the forms and aspects of thinking, which are mainly used by teachers to diversify students. A document analysis guide is needed to help researchers analyze document data.

2. Interview

This research used semi-structured interviews. In the process, the researchers prepared and asked questions based on the research questions. The interview aimed to investigate that how if the English teachers use higher-order thinking skills assessments in their test for their students. The interview guide is required to help the researcher conduct the interview.

3. Test

The research process is research and development, the researchers need test questions used by teachers so that they can be developed into types of test questions that fulfill the HOTS standards. The purpose of knowing the types of previous questions also helps researchers to analyze the form of previous test so that they can be developed as much as possible. This also helps researchers to see the media and types of questions used by teachers so that in collecting data, researchers will find it easier to collect.

4. Validation Checklist

Validation checklist was used as a benchmark and given to both of expert in evaluating the product. This validation checklist used to measure the level of validity of each items.

3) RESULTS

The Existing of Test Instrument

Documents, interviews, and teacher's test items are the components to be analyzed at this section. In this case, the document consists of the Syllabus, Lesson Plan, and teaching materials.

1. Document analysis

- a. Syllabus

Language learning cannot be separated from the requirements of learning such as how to plan, how to design, and how to obtain learning resources. All of the above are closely related to the syllabus. The syllabus contains a framework or details of how teachers and students should act in the classroom. The syllabus is also a guide or instructions for organizing activities in the classroom.

The available syllabus was in accordance with the provisions of the 2013 curriculum because it already consists of several components that must exist in the development of the

syllabus, including: competency standards, basic competencies, learning materials, indicators, assessments, time allocation, and learning resources developed by each educational unit.

b. Lesson Plan

Lesson plan mandatory for every subject teacher in every education unit, including the X grade English teacher at MAN Madani Pao-Pao. RPP is described from the syllabus to direct student learning activities in an effort to achieve KD. Every teacher in the education unit is obliged to prepare a complete and systematic lesson plan so that learning takes place interactively, inspiring, fun, challenging, motivating students to participate actively, and providing sufficient space for initiative, creativity, and independence according to talent, interests, and the physical and psychological development of students.

c. Textbook

Textbook is primary sources that contain learning materials for achieving basic and core competencies. Competence is intended to be related to increasing faith, piety, noble character and personality, mastery of science and technology, increasing sensitivity and aesthetic ability, increasing kinesthetic ability, and improving health, as measured by national education standards. The National Education Standards Agency (BNSP) has established that an excellent book is one that meets four criteria for eligibility: material feasibility, linguistic feasibility, presentation feasibility, and graphic feasibility.

2. Interview analysis

The interview technique was used by determining the representative sources in line with the research questions. The responders in this study were English teachers, specifically Mrs. Muslihah. On May 22, 2021, the interview was held at his home in Paccinongan.

Researchers interviewed respondents to learn about the teacher's comprehension of HOTS while developing questions. The results of the interview above:

“Both of them. Because of the first one, Essays are easy to measure, after all, students' abilities. If the multiple choice is determined by the school”

Next, whether or not to include opening texts, pictures, and scenarios while creating questions? She responded:

“There are introductory texts, visuals. Oh yes, there are.. there are.. a lot because it's like matching the picture with what is the description of the picture. What is the description of a picture, Describe something. And scenarios also exist”

In response to further questioning about the aim of implementing these test items, she stated:

“That definitely measures whether he understands what he's been learning all this time or not”

3. Teacher's Test Items Analysis

The following test was given by the teacher to students through the Google form along with analysis from researchers which will then be developed by researcher. The researcher used the cognitive level theory to assess the questions posed by the teacher. Anderson and Krathwohl (2001) categorize thought process dimensions as follows:

Table 1. Thought Process Dimension

LOTS (Lower Order Thinking Skill)	C1 (Remembering)	<ul style="list-style-type: none"> • Recalling. • Verbs: remember, register, repeat, imitate.
---	------------------	---

MOTS (Middle Order Thinking Skill)	C2 (Understanding)	<ul style="list-style-type: none"> • Explain ideas/concepts. • Verbs: explain, classify, accept, report.
	C3 (Applying)	<ul style="list-style-type: none"> • Using information on a different domain • Verbs: use, demonstrate, illustrate, operate
	C4 (Analyzing)	<ul style="list-style-type: none"> • Specify aspects/elements • Verbs: compare, examine, criticize, test.
HOTS (Higher Order Thinking Skill)	C5 (Evaluating)	<ul style="list-style-type: none"> • Take your own decisions • Verbs: evaluate, judge, refute, decide, choose, support
	C6 (Creating)	<ul style="list-style-type: none"> • Create your own ideas/ideas • Verbs: construct, design, create, develop, write, formulate

Developing of HOTS Test Instrument

Plump development model is used in the construction of the test item. This model has five stages of development: (1) preliminary study, (2) design, (3) realization/construction, (4) test, evaluation, and revision, and (5) implementation (Implementation). This study's focus on generating HOTS questions includes a grid of questions, questions, and grading standards (rubrics). The development of items pertains to the validity and dependability requirements.

1. Preliminary Investigation Phase

The main activity at this stage is analysis the document-document for the development of new product and analysis the feasibility and requirements of developing of HOTS test items. Therefore, the researchers conducted preliminary observations based on document analysis and interviews analysis with the 10th grade English teacher of MA Madani Alauddin Pao-Pao.

2. Design Phase

a. Collecting

To get to this point, various papers must be prepared, including the syllabus, lesson plans, teaching materials, and teacher-created questions. The researcher evaluated these materials as part of the process of creating the questions into HOTS questions.

b. Organizing

Several findings of the study of these documents were derived from the data collection acquired. The syllabus is utilized as a guideline while creating a question grid, which is also used to produce the questions to be created. In addition to the syllabus, teaching materials serve as a guideline when developing questions. This teaching material is divided into ten chapters, with the questions provided by the researcher focusing on the chapter "Why Were They Famous?" This chapter focuses on reading skills that discuss about recount text, particularly in biographies. Several biographies of prominent people are offered in this content, together with their literary structures. There are also some quizzes and vocabulary concerning the biography itself. Based on these materials, the content of the questions must be in sync with what students have learnt about this material while generating the substance of the questions.

c. Creating

The next stage is to create after all of the documents have been organized. At this stage, files that did not previously exist were created to assist the production of a question product. For example, lattices, cover questions, item grading rules, and expert validation tools.

3. Construction/Realization

At this step, a prototype, namely the primary design based on the initial concept, is created. In the context of education, the second and third steps mentioned above are referred to as the production stage. It means, the HOTS test instrument is created. There are 10 numbers of multiple choice and 5 numbers of Essay. The researchers develop the test being multiple choice and essay.

Based on the findings of the analysis of the teacher's questions, the researcher created questions based on the HOTS criteria by Alderson.

4. Test, Evaluation and Revision Phase

This stage is concerned with the quality of the design that will be created. Make judgments after deep reflection as well. The process of gathering, processing, and evaluating information in a systematic manner is referred to as evaluation. The first step at this stage was all of the items were validated by experts before the items were used. The suggestions were given by the experts can be seen below:

Table 2. Revision from Expert

Revision from Expert I (Indah Miftah Awaliyah S.S., M.Hum)
The question narration should be more clear
The question narration don't use (what, when, and who)
Use a diagram, picture and etc
The picture that used proper to modern view

Revision from Expert II (Annisa Shofa Tsuraya S.Pd., M.Pd)
Put in a shape each passage and attach the picture of the biography
Pay attention on the use of punctuation.
Pay attention for grammatical rules
Used HOTS vocabulary (create, analyze)

After knowing what was will be revised, the researcher re-make the product based on the revision. The new product, namely Prototype II. The product/prototype II was the test instrument based on the HOTS concept with revision by the expert. The new product can be seen in this link: https://drive.google.com/drive/folders/1aDzRGgeOPHte9xksl-px_Zuc4dSamKT?usp=sharing

The classification of the product above based on cognitive level can be seen as follows:

Table 3: Analysis of Multiple-Choice Test Instrument

Test Items	Cognitive Level	Category
1 st Question	C4 (Analyze)	HOTS
2 nd Question	C4 (Analyze)	HOTS
3 rd Question	C4 (Analyze)	HOTS
4 th Question	C5 (Evaluating)	HOTS
5 th Question	C4 (Analyze)	HOTS
6 th Question	C4 (Analyze)	HOTS
7 th Question	C5 (Evaluating)	HOTS
8 th Question	C4 (Analyze)	HOTS
9 th Question	C4 (Analyze)	HOTS
10 th Question	C4 (Analyze)	HOTS

Table 4: Analysis of Essay Test Instrument

Test Items	Cognitive Level	Category
1 st Question	C6 (Creating)	HOTS
2 nd Question	C4 (Analyze)	HOTS
3 rd Question	C4 (Analyze)	HOTS
4 th Question	C4 (Analyze)	HOTS
5 th Question	C6 (Creating)	HOTS

After the new product was created, the next phase was test trial to the students. It's aims for measuring the empirical validity and reliability of the product.

5. Implementation Phase

After analyzing the items developed, the quality of the items was obtained which will be implemented by the English teacher of class X MIA 2 MA Madani Alauddin Pao-Pao with the address at Paccinongan, Somba Opu, Gowa, South Sulawesi. At this stage, it will be implemented in schools as a bank question on HOTS.

The Level of Validity and Reliability of Test Items

1. Expert Validity

Table 5. The result of Validation sheet by expert

Nomor Item	Aspek yang Dinilai	Penilaian Validator	
		Validation 1	Validation 2
MATERI SOAL YANG DISAJIKAN			
1.	Keterkaitan antara soal dengan indikator soal	3	4
2.	Keterkaitan soal dengan tingkat perkembangan intelektual peserta didik	3	4
3.	Soal bervariasi sesuai dengan tingkatan kognitif	3	4
	Rata-rata	3	5,3

KONSTRUKSI			
1.	Pertanyaan pada soal dinyatakan dengan jelas	3	4
	Rata-rata	3	4
BAHASA			
1.	Menggunakan kalimat yang komunikatif	3	3
2.	Menggunakan bahasa yang sesuai dengan kaidah bahasan Indonesia yang benar	3	3
3.	Menggunakan bahasa yang sederhana dan mudah dimengerti	3	3
4.	Kalimat dalam setiap item soal tidak menimbulkan penafsiran ganda	3	3
	Rata-rata	3	3
SAMPUL BANK SOAL			
1	Menggunakan desain sampul yang menarik	3	3
	Rata-rata	3	3

2. Empirical Validity

The analysis of the empirical validity and reliability of the HOTS items that have been developed can be seen from the test of the validity of the questions using correlation by *Statistical Product and Service Solution* (SPSS) program. Empirical validity testing, the researcher immediately tried it out instrument to students and then the results are analyzed with the validity test by using of the SPSS program. Based on the results of the students' work, the level of test validity can be calculated. The following data is the result of the calculation of the validity of the test based on SPSS.

Table 6. The Result of Multiple Choices and Essays validity

No	Nilai r_{hitung}	Nilai r_{tabel}	Kategori
1	0,647	0,444	Valid
2	0,606	0,444	Valid
3	0,695	0,444	Valid
4	0,657	0,444	Valid
5	0,619	0,444	Valid
6	0,604	0,444	Valid
7	0,735	0,444	Valid
8	0,581	0,444	Valid
9	0,736	0,444	Valid
10	0,542	0,444	Valid
No	Nilai r_{hitung}	Nilai r_{tabel}	Kategori
1	0,792	0,444	Valid

2	0,649	0,444	Valid
3	0,628	0,444	Valid
4	0,623	0,444	Valid
5	0,696	0,444	Valid

3. Reliability Analysis

This reliability test is based on the results of trials validation test previously which all the test items are valid. So that the level of test reliability can be calculated by using SPSS application.

a. Multiple Choice

Table 7. The Result of Multiple Choices Reliability

Case Processing Summary

		N	%
Cases	Valid	20	100.0
	Excluded ^a	0	.0
	Total	20	100.0

Reliability Statistics

Cronbach's Alpha	N of Items
.841	10

b. Essay

Table 8. The Result of Essays Reliability

Case Processing Summary

		N	%
Cases	Valid	20	100.0
	Excluded ^a	0	.0
	Total	20	100.0

Reliability Statistics

Cronbach's Alpha	N of Items
.723	5

4) DISCUSSION

Based on the findings, there were three aspects that need to be summarized. The first was the existing of the test instruments. Both of syllabus and lesson plan were not suitable with 2013 curriculums, but the textbook was suitable with 2013 curriculum. The result of interview analysis indicated that the teacher had been having one of the criteria for HOTS criteria, namely the use of a variety of questions in the creation of items, multiple choice, and essays. It is generally seen that the number of questions she made for the final exam per semester is 35 numbers, where each chapter represents 7 numbers, but depending on the number of chapters, she divides the number of numbers as evenly as possible for each chapter. Furthermore, for the questions, she created her own, although not far from the textbook, and some pictures were taken from the textbook.

There were 6 teacher's test items, the first question can be classified into low level based on the HOTS (Higher Order Thinking Skill) concept because the question just applied remembering skill (C1). It didn't need analysis skill and comprehension skill which these two-skill classified in HOTS level. Related with the first question, the second question can be classified into low level, because it's just needed a remembering skill (C1), The students just need memorizing thi material about the verb 2. The third question and the fourth question also can be classified into low level but the fifth and the sixth question classified into middle level which need applying skill (C3). didn't need a critical thinking for answering that.

Based on the findings of the analysis of the teacher's questions, the researcher created questions based on the HOTS criteria, there were 10 multiple choice and 5 essays. The first question consist a biography of Ir. Soekarno, which was well-known among schoolchildren. Because the presence of an introduction or stimulus in the form of things that pupils relate to was one of the features of HOTS questions. Questions about statements not covered in the reading for question number one. The second query concerns the purpose of reading. The final question was regarding the word equivalents based on the reading. These three questions were included in the HOTS category because in order to answer them, the reader must first thoroughly analyse the text, or in other words, pupils must rely on their analytical skills. Similarly, practically all multiple-choice problems rely on analytical skills for the following number. Some essay questions required students to write a biography or a summary. Whereas queries like this were into the HOTS category, namely 'Creating.' In addition, some of them feature queries that require the capacity to analyse and recognize.

These test items had revised by the expert and the researcher made a new product based on the suggestion by expert. After the new product was created, the next phase was test trial to the students. It's aims for measuring the empirical validity and reliability of the product.

The result of validation by expert that was developed in this study, it is in a score of 3.7 which is very valid category, based on predetermined criteria. HOTS items with aspects of the material then presented with a value of 3.5 are in the veryvalid category, the construction aspect shows a value of 4 in the very valid category, then the language aspect shows a value of 3.5 which is in the very valid category. The higher the VR (close to 4 or equal to 4), the higher the validity value of an item/item, and the lower the VR number (close to 0 or equal to 0), the lower of validity value of an item.

The result of empirical validity, the analysis of the empirical validity and reliability of the HOTS items that have been developed can be seen from the test of the validity of the questions using correlation by *Statistical Product and Service Solution* (SPSS) program. Technique that will be used by researchers is the value of 5% significance with a critical value. In other words, can compare between rhitung with rtabel. Test results the validity of the data shows that it is valid if r-hitung is larger than r-tabel. The next, the reliability analysis. The level of test reliability can be calculated by using SPSS application. Meanwhile, the tool that is used to get the realibility in this

research by using Alpha Cornbach, that is SPSS program. The criteria of the reliability test is if the coefficient Alpha is larger than the significance 0,6, then the test are reliable. Based on the data above, the result of reliability in multiple choice test by Cronbach's alpha is 0,841 which means all the test items are reliable. Beside, the essay test instrument shows score in 0,723 which means all the essay items test are reliable.

Furthermore, in order to be effective, the test should not only be valid and accurate, but it should also follow some criteria. The researcher mentioned certain language assessment principles. According to Brown (2003:19), an effective test must include some language assessment standards. There are several types of washback: practicality, reliability, validity, authenticity, and washback. The following are the test instruments based on the findings of this study; Practicality. The student does not require a large sum of money based on the results of the test instrument. In this situation, the test is not prohibitively expensive. The form of the test was created by the researcher using a Google form that was created online. The test does not need to be printed by the teacher. The test then does not require much time to complete. The test may then be administered conveniently and automatically in terms of scoring and grading. The next, reliability. The test is reliable and consistent. Using the SPSS tool, the researcher examined the reliability of the results of the respondents' answers. As a result, the Cronbach's alpha result for reliability in multiple choice tests is 0.841, indicating that all test items are reliable. Furthermore, the essay test instrument yields a score of 0.723, indicating that all of the essay items tested are reliable.

The third is validity. According to the SPSS program's results, the 10 multiple-choice questions and 5 essay questions are declared valid because all of the calculated r -hitung are more than the r -table values. With a total of 20 student respondents, it is declared that the 10 multiple choice questions assessed and 5 essays are valid. Next, Authenticity. It is a test that is based on the evaluation or construction of the test. The test instrument was created using the lesson plan, syllabus, and material. Furthermore, the test is based on the HOTS category and has been evaluated by ELT teachers. And the last Washback. The final option is washback. The test instrument is being returned to the students. The final test was reviewed and confirmed while being built. The test was carried out by the researcher using the HOTS category. In this situation, the test referred to HOTS characteristics as defined by Jurnal Penyusunan Soal HOTS (2017). It is possible to conclude that the test is washback since the test can cause pupils to think more critically while analysing the test.

5) CONCLUSION

There were two limitations of this research, response and implemented. Obtaining data from students who wanted to be examined proved problematic for researchers during this study. According to the time when the test questions were disseminated via the Google form, relatively few students answered. This hurdle, though, can be surmounted in the end. The only stage of this study that was not carried out was the implementation stage. As a result, this stage will be carried out by the teacher in question. The researcher's product will be directly implemented by the teacher on the pupils who will appear during the final exam later.

REFERENCES

- Abosalem, Yausef. (2016). *Assessment Techniques and Students' Higher-Order Thinking Skills*. International Journal of Secondary Education. Vol. 4(1). 1-11.
- Anderson, L.W. and D.R. Krathwohl. 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision on Bloom's Taxonomy of Educational Objectives*. New York: Longman
- Anggreani, 2017. *Analysis Of Learning Implementation Plan (Rpp) Indonesian Teacher Sma Negeri 7 Medan Year Of Study 2016/2017*.

- Arikunto, S. (2005). *Dasar-dasar evaluasi pendidikan*. Jakarta: Bumi Aksara.
- Arikunto, Subarsini. *Prosedur Penelitian Suatu Pendekatan Praktek*. Jakarta : Rineka Cipta. 2002
- Bloom, Benjamin S. *Taxonomy of Educational Objectives*. Ann Arbor: David McKay Company Inc, 1956.
- Brown, D. 2004. *Language Assessment: Principles and Classroom Practices*. New York: Longman
- Chakanyuka, S. (2000). *Measurement and evaluation*. Harare: Zimbabwe Open University.
- Djiwandono, Soenardi. 2011. *TES BAHASA Pegangan bagi Pengajar Bahasa*. Malang Edisi 2: PT Indeks
- Febrina, Usman, B., Muslem, A. 2019. *Analysis of Reading Comprehension of Questions by Using Revised Bloom's Taxonomy on Higher Order Thinking Skill (HOTS)*. English Education Journal. Vol 1(1). Hal 1-5.
- Harefa, A. O. (2009). *Penilaian dan Hasil Belajar. Didaktik: Jurnal Ilmiah Pendidikan, Humaniora, Sains, dan Pembelajarannya*, 3(1), 15-31.
- Mardapi, D., Kumaidi. & Kartowagiran, B. (2011). *Pengembangan Instrumen Pengukur Hasil Belajar Nirbias dan Terskala Baru*. *Jurnal Penelitian dan Evaluasi Pendidikan* 15(2): 326-341
- Mirnasulistiyawati, 2018. *Analisis Penerapan Kurikulum 2013 Dalam Meningkatkan Kualitas Pembelajaran*. Fakultas Tarbiyah Prodi MPI Institut Agama Islam Negeri Bone, Sawange.
- Muhsin Arief M. (2015). *The Correlation Between Students' Grammar Knowledge And Writing Ability*. (skripsi Muhammadiyah University of Makassar). Makassar
- Narwianta, N., Bharati, L, A, D., & Rukmini, D. 2019. *The Realization of Higher Order Thinking Skills in English School Nationally Standardized Examination at State Senior High School 6 Semarang*.
- Nugraha, Alga. 2019. *Pengembangan Instrumen Penilaian Kognitif Kemampuan Berpikir Kreatif PJOK pada Materi Permanen Sepak Bola Peserta Didik SMA*. (Tesis, Universitas Negeri Yogyakarta). Yogyakarta
- Nugraha, Alga. 2019. *Pengembangan Instrumen Penilaian Kognitif Kemampuan Berpikir Kreatif PJOK pada Materi Permanen Sepak Bola Peserta Didik SMA*. (Tesis, Universitas Negeri Yogyakarta). Yogyakarta
- Nurwanah. 2019. *Pengembangan Butir Soal Kemampuan Berpikir Tingkat Tinggi pada Mata Pelajaran BIOLOGI Kelas XI SMA Negeri 3 Pangkep*. (Skripsi, UIN Alauddin Makassar). Makassar.
- Peterson, D., Kromrey, J., Borg, J., & Lewis, A. (1990). *Defining and Establishing Relationships between Essential and Higher Order Teaching Skills*. *The Journal of Educational Research*, 84(1),5-12. Retrieved from <http://www.jstor.org/stable/40539677>.
- Putra, T. K., & Abdullah, D. F.: *Higher-order thinking skill (HOTS) questions in English national examination in Indonesia*. Vol.11(1), pp. 145-160. *Jurnal Bahasa Lingua Scientia*, (2019)
- Rochmad (2012). *Desain Model Pengembangan Perangkat Pembelajaran Matematika* *JURNAL KREANO*, ISSN: 2086-2334 Volume 3 Nomor 1, Juni 2012 Jurusan Matematika FMIPA UNNES
- Wayan Widana, I. *Higher Order Thinking Assessment (HOTS)*. Volume 3. No. 1 JISAE Ikacana Publisher. 2017
- Westwood, P. (2008). *What teachers need to know about teaching methods*. Victoria: Acer Press
- Widana, I Wayan. 2017. "Higher Order Thinking Skills Assessment (HOTS)". *Journal of Indonesian Student Assessment and Evaluation*. Vol 3 No.2, 2017. 32
- Yassi dan Kaharuddin. 2018. *Syllabus Design for English Language Teaching*. Jakarta: Edisi 2: Prenadamedia Group.