THE EVALUATION OF HIGHER ORDER THINKING SKILLS IN ENGLISH SCHOOL NATIONALLY STANDARDIZED EXAMINATION AT *STATE SENIOR HIGH SCHOOL 6 SEMARANG*

A Thesis

Submitted in partial fulfillment of the requirements for the degree of Magister Pendidikan in English Language Education

by
Nanang Narwianta
0203517074

ENGLISH LANGUAGE EDUCATION
PASCASARJANA
UNIVERSITAS NEGERI SEMARANG

APPROVAL

This thesis entitled **"The Evaluation of Higher Order Thinking Skills in English School Nationally Standardized Examination in State Senior High School 6 Semarang"** written by,

Name : Nanang Narwianta

NIM : 0203517074

Department : English Language Education

has been examined and defended on July 15th, 2019 in front of the Board of Thesis Examiners.

Semarang, July 2019

The Board of Examiners

Chairman,

Prof. Dr. Totok Sumaryanto F, M.Pd
NIP. 196410271991021001

Secretary,

Prof. Dr. Abdurrachman Faridi, M.Pd.
NIP. 195301121990021001

First Examiner,

Prof. Dr. Januarius Mujiyanto, M.Hum.
NIP. 195312131983031002

Second Examiner,

Prof. Dr. Dwi Rukmini, M.Pd
NIP.195104151976032001

Third Examiner,

Dr. Dwi Anggani Linggar Bharati, M.Pd
NIP. 195901141989012001

PERNYATAAN

Dengan ini, saya

Nama : Nanang Narwianta

NIM : 0203517074

Program Studi : Pendidikan Bahasa Inggris

menyatakan bahwa yang tertulis dalam thesis yang berjudul berjudul "The Evaluation of Higher Order Thinking Skills in English School Nationally Standardized Examination at State Senior High School 6 Semarang" ini benar-benar karya saya sendiri bukan jiplakan dari karya orang lain atau pengutipan dengan cara-cara yang tidak sesuai dengan etika keilmuan yang berlaku baik sebagian atau seluruhnya. Pendapat atau temuan orang atau pihak lain yang terdapat dalam disertasi ini telah dikutip atau dirujuk berdasarkan kode etik ilmiah. Atas pernyataan ini, saya **secara pribadi** siap menanggung resiko/sanksi hukum yang dijatuhkan apabila ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya ini.

Semarang, 15 Juli 2019

Yang membuat pernyataan,

Nanang Narwianta

0203517074

MOTTO AND DEDICATION

Motto

Ever tried. Ever failed. No matter. Try again. Fail again. Fail Better

Samuel Beckett

Dedication:

For the Bangka Belitung Province Government, SMA Negeri 1 Membalong and The Civitas Academica of Universitas Negeri Semarang

ACKNOWLEDGEMENT

Alhamdulillahirobil`alamin. All the praises be to Allah. First of all , I would like to take opportunity to express my gretest gratitude to Allah SWT, for His blessing and guidance so that this thesis finished well.

My sincerest thankfulness goes to my supervisors, Dr. Dwi Anggani Linggar Bharati, M.Pd and Prof. Dr. Dwi Rukmini, M.Pd for their patience, helpful corrections, and positive criticism during my consultations for this thesis.

The writer also would like to express the thanks to Prof. Dr. H. Achmad Slamet, M.Si. as the director of Pascasarjana, Universitas Negeri Semarang, Prof. Dr. Totok Sumaryanto Florentinus, M.Pd as the first deputy director, and Prof. Dr. Tri Joko Raharjo, M.Pd as the second deputy director of Pascasarjana, have allowed the writer to study in Pascasarjana , Universitas Negeri Semarang.

My deep gratitude is extended to Drs. Bambang Widi Pratolo, M.Hum. Ph.D as my expert investigator of the instrument and the data, who has given the criticisms, comment and supporting advice for the improvement of this report.

My honor also goes to all lecturers of the English Departement of Universitas Negeri Semarang for all the guidance and lectures during my study. I also would like to thank to Dra. Lukita Yuniati, M.Kom,the principal of SMA Negeri 6 Semarang and the English teachers of that school, especially Dra. Ida Rahmawati and Dwi Budi M.Pd for the permission and cooperation during this research.

The biggest thanks are dedicated to the Harjo Mulyono`s family for their endless love, supports and prayers. Special thanks are devoted to my lovely wife;

Nisfi Rahmawati and my little boy Nafian Eshaal Alindra. My boarding house partner Rizky Hendrawan and all my friends at Rombel 3 and 4 that cannot state one by one and all people for encouragement, great support, and greater sense of togetherness.

I do realize that there are still many shortcoming in this thesis. Therefore, I would be grateful for any criticisms and suggestions for its betterment.

ABSTRACT

Narwianta, Nanang. (2019). *The Evaluation of Higher Order Thinking Skills in English School Nationally Standardized Examination at State Senior High School 6 Semarang*. Thesis, English Language Education Pascasarjana Universitas Negeri Semarang. Supervisor I. Dr. Dwi Anggani Linggar Bharati, M.Pd. Supervisor II. Prof. Dr. Dwi Rukmini, M.Pd

Keywords: Evaluation, HOTS, English Nationally Standardized Examination

The English school nationally standardized examination as an assessment of the learning outcomes conduct by educational unit is designed to measure the student`s achievement competencies, to determine student`s graduation from the school, it is very significant to conduct test items evaluation since it gives a clear portrait of the quality of the items and the test as a whole.

The purpose of this study was to explain the realization of higher order thinking skills reflected in English school nationally standardized examination at state senior high school 6 Semarang academic year of 2018/2019. Then this research evaluate whether the items in English school nationally standardized examination met the requirement of content validity, reliability and practicality.

The researcher use the mixed method using analysis card to analyze the realization of higher order thinking skills in English school nationally standardized examination. The content validity analyze through matching method between the basic competensies and the question in paper test.While to ensure the reliability of higher order thinking skills in English school nationally standardized examination, use the KR-21 for multiple choice test items and alpha cronbach for essay test items. And the practiality obtained through checklist instrument.

The results indicated that HOTS realized in listening, reading, and writing questions. The listening in the form of spoken written, because it is to test listening simultaneously with speaking. There are one listening question categorized into HOTS of analyzing level, in reading the questions belonging to HOTS reach 8 questions that consist of 5 questions of analyzing level and 3 questions of evaluating level. Meanwhile, there is one question of HOTS into creating level in writing questions. The results also showed the proportion of HOTS comprised 22,22% met the requirement from BNSP that stated in academic year 2018/2019 there is approximately 10-15% of higher order thinking skills test items. Meanwhile, the content validity of the questions is high, and then the reliability of those test items categorized as unreliable and the practicality meet the requirements of the practical test items. Accordingly, this finding reveals that there is still much room for the variation of HOTS levels implemented all skills in English school nationally standardized examination. In the light of these result, this study recommends modifying the proportion of HOTS in four skills by providing them with more representation of the content validity.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

LIST OF APPENDICES

Appendix **Page**

CHAPTER I

INTRODUCTION

This introduction includes the background of the study, reasons for choosing the topic, research problems, objectives of the study, significance of the study, and the outline of the thesis.

1.1 Background of the Study

Curriculum 2013 has been launched by the government since the academic year 2013/2014 and implemented gradually at all levels of school throughout Indonesia. The curriculum 2013 is prepared for generations to face the globalization. The government develops the quality of education by enhancing student potential for competitiveness in the 21st century. In this case, the curriculum 2013 focuses on literacy, 4Cs (critical thinking, creative, communication, and collaboration), and characters. The main objective of the curriculum 2013 is equipped graduates to be productive, creative, innovative, and affective through strengthening integrated attitudes, knowledge, and skills.

The English subject is part of the curriculum 2013, which the scope of administration starts from formulating basic competencies, attitude competencies developed in English subjects indirectly and integrated into the learning process. English subjects for senior high school develop the potential of students to have communicative competence in interpersonal, transactional, and functional texts, using a variety of spoken and written English texts. Through the use of these texts, students are guided to use

factual, conceptual, and procedural knowledge, and to encourage noble values of national character, in the context of life in the home, school, and society.

Teaching English in senior high school is focused on increasing the competency of students to use the language to achieve communication goals in various contexts, both spoken and written with higher complexity of the material studied in the previous level, using the same approach, namely text-based approach. This learning refers to the function of language and its use, which is a unity of meaning both spoken and written. In addition, the text-based approach also emphasizes increasing the ability of students to use English in various types of texts. The text is studied not as a final goal, but as a tool for carrying out various activities in real life. At this intermediate level, teaching material consists of simple texts.

In general, English competencies that must be mastered by senior high school graduates is the ability to communicate in three types of texts, (1) interpersonal, (2) transactional, and (3) functional, spoken and written, at the informational literacy level, to carry out social functions, in the context of personal, socio-cultural, academic and professional, using various forms of text, with coherently and cohesively acceptable structures and appropriate linguistic elements.

The aspect that also changes in curriculum 2013 is the assessment. Frankly, assessment in curriculum 2013 become a specific problem. Wijayanti (2015) stated that English teachers were not ready to implement of assessment in curriculum 2013. Assessment is an important component in the

education system to determine the progress and level of achievement of learning outcomes. It is done to make sure whether the learning process has been running well over the term. The objective of the assessment itself is to help the teacher ascertain the degree to which educational objectives have been achieved, to review the effectiveness of teaching method and to help the teacher know his students as an individual.

Recently the implementation of higher order thinking skills in the learning process also happens in the assessment. Assessment that oriented higher order thinking skills is not a new form of assessment for the teacher in an assessment program. But this higher order thinking skills oriented assessment maximizes the teacher's skills in making judgments. Teachers in this assessment must emphasize the assessment of attitudes, knowledge, and skills that can improve students' skills in the higher order thinking skills-oriented learning process.

Assessment of learning outcomes by education units is conducted to measure student competencies as recognition of learning achievement of the education unit conducted through school nationally standardized examination or in Indonesian abbreviation USBN is the final test conducted by teachers and the education unit. It is an internal assessment to measure what have students learned. Because the school nationally standardized examination results determine graduation from education units, the test items are expected to meet the requirements good instrument to provide valid and objective information. Test items that are not good provide information that is not in

accordance with the achievements students so that they can harm students and provide incorrect information or misleading decision-makers.

The English school nationally standardized examination covers four skills, they are listening, speaking, reading and writing. In this case, the listening and speaking skills assess integrated through the spoken written form then reading and writing questions in written form. In English school nationally standardized examination speaking assesses integrated with listening questions. Here students are required to give the response to the questions, the students answer the question not orally but choosing the correct answer in written form. So, the speaking skill assesses through the spoken written form simultaneously with the listening questions because in senior high school the degree of literacy at the informational level, it means students are able to access knowledge with language skills. It is in line with Widdowson (1996, p.59) says that the skill of speaking involves both receptive and productive participation. The receptive aspect of speaking is a skill which is conventionally referred to as listening while the productive aspect of speaking referred to as saying. It can be said that speaking has a productive part when one participant in an interaction assumes the active role of speakers.

Since the academic year 2016/2017, education unit tests on several subjects upgraded to School Nationally Standard Examination for junior high school, senior high school, and vocational school. In some subjects, 25% of test items provided from the Ministry of Education and Culture, and 75% of

the test items are constructed by teachers through teacher forum (MGMP). Writing about school nationally standardized examination test items is important because it was written by the teacher in each education unit.

To assess students' higher order thinking skills, questions are designed that students answer the questions through thinking processes in line with operational verbs in Bloom's taxonomy, both in terms of knowledge, attitudes, and skills. In learning especially English, it is stated that the ability of students is not only to master a set of knowledge in the form of facts, concepts, or principles but also a process of discovery; it means students must always be invited to learn with using the thought process to find out these concepts.

It is important to learn higher-order thinking skills. Brookhart (2010) describes the type of higher order thinking skills based on the purpose of learning in class, which consists of three categories: higher order thinking skills as a transfer of knowledge, higher order thinking skills as critical thinking, and higher order thinking skills as problem-solving. Higher order thinking skills as a transfer of knowledge are defined as the skill to apply the knowledge and skills that have been developed in learning in a new context. The new context is interpreted as something that has not been taught before. Higher order thinking skills as a transfer of knowledge includes analyzing, evaluating, creating, creative thinking, and logical thinking by summarizing it to analyze, evaluate and create, included creative and logical thinking. Higher order thinking skills as critical thinking are defined as judgment skills using

logical and scientific reasons. This includes critical and metacognitive thinking. Higher order thinking skills as problem-solving is defined as the skill of identifying problems and resolving non-structured problems. This includes problem-solving itself. So, the higher order thinking skills include analyzing, evaluating, creating, critical thinking and problem-solving. Indicators of analyzing, evaluating and creating skills are based on the theory presented by Anderson & Krathwohl (2001, p.28), while critical thinking skills and problem-solving are based on the theory described by Brookhart (2010, p.98).

Higher order thinking skill as explained by Thomas & Thorne (2009) is a thinking skill more than just memorizing facts or concepts. Higher order thinking skills require students to do something about these facts. Students must understand it, analyze, categorize, manipulate, create new ways creatively and apply them in finding solutions to new problems. To carry out the assessment, the teacher needs an assessment instrument in the form of questions, to assess the aspects of knowledge, attitudes, and skills. The assessment instrument used by the teacher to assess the learning outcomes of students on aspects of knowledge is usually taken from various books or a collection of examination questions. Questions can be in the form of an essay or multiple choices.

The background of the promotion of higher order thinking skills items is the lacks English skills of Indonesian students in surveys conducted by international benchmarking such as PISA, TIMMS, and PIRLS. Learning

critical thinking is not straightforward, such as learning about the material, but learning how to think critically to solve problems is related to one another. Thinking skills of students can be trained through activities by given the students a problem, in this case, the variation problem in the form of the questions. The reality in the field shows that English questions tend to test more aspects of memory. Many books present material by inviting students to active learning, the presentation of concepts is very systematic, but often ends with questions of evaluation that lack training in high order thinking skills of students.

High order thinking skills should be implemented and accustomed to English skills. According to the core competencies for senior high schools, learning is designed to provide experience in using English texts to understand and apply factual, conceptual, and procedural knowledge related to real situations and events, through listening, speaking, reading and writing in concrete and abstract domains. Listening means understanding various meanings (among individuals, opinions of textbooks) various oral texts that have communicative goals, text structures, and certain linguistics. Speaking means expressing various meanings (among individuals, opinions, textbooks) through various oral texts that have communicative goals, text structures, and linguistic features. Reading means understanding various meanings (among individuals, opinions, textbooks) in various written texts that have communicative goals, text structures, and certain linguistic features. Writing means revealing various meanings (among individuals, opinions, textbooks)

in various written texts that have communicative goals, text structures, and linguistic features. The use of text also aims to foster an attitude of respect and appreciate religious and social values, including honesty, discipline, responsibility, caring, kinds, and confidence, in interacting effectively with the social and natural environment within social contact and its existence.

Based on the curriculum 2013 the purpose of learning English can be achieved by implementing higher order thinking skills. It is in line with the government that expects students to achieve various competencies by applying higher order thinking skills. These competencies are creative and innovative, communication skill, collaboration, and confidence. Higher order thinking skill is defined as the level of thinking where the process is done more than just repeating information or facts. Thomas & Thorne (2009) explain that higher order thinking skills as a way of thinking at a higher level than memorizing or retelling something that is told by other people. Furthermore, King, Goodson, & Rohani (2013) argue that higher-order thinking skills are the ability to think that not only requires the ability to remember, but also higher capabilities. Higher order thinking skills are student`s abilities that are activated when students encounter unfamiliar problems, uncertainties, questions or dilemmas. Whereas in Bloom's taxonomy revised by Anderson & Krathwohl (2001, p.28), higher order thinking skills are specified in three-dimensional thinking consist of analyzing, evaluating and creating.

Higher order thinking skills should be implemented in the learning process and assessment program. Students should be trained and accustomed to higher order thinking skills during the learning process and an assessment program like the formative and summative test. Higher order thinking skills certainly not effective if it is not preceded in the learning process. Assessment of higher order thinking skills in curriculum 2013 is needed as an effort to improve the quality of learning and improve the quality of graduates. National examination and school nationally standardized examination as the summative test is regarded as the high-level assessment program but because of the limitation of the time and budget not all of higher order thinking skills can be covered in school nationally standardized examination test items.

In constructing school nationally standardized examination test item, there is a possibility of teacher-made questions of the cognitive level lower than expected in the blueprint. Generally, school nationally standardized examination test items are constructed by teachers only measure the lower level of thinking skills. Other causes are not inserted higher order thinking skills in the school nationally standardized examination that cause learners not accustomed to working on higher order thinking skills test item. On the other hand, in nationally examination test items learners are required the ability to work on higher order thinking skills questions. Every year the percentage of higher order thinking skills test items which are inserted in the national examination is improved. Board of Education National Standard or BSNP (2018) stated that in the national examination year 2018/2019 there are

approximately 10-15% of higher order thinking skills test items. Therefore, higher order thinking skills should be prepared and inserted in school nationally standardized examination test items.

Concerning the problems that the writer found on the preliminary study, the writer believes that English school nationally standardized examination is very important to determine the students` mastery of higher order thinking skills. Besides, the outcome of school nationally standardized examination become one of the element to determine students` graduation from the educational unit. Higher order thinking skills of item test on school nationally standardized examination is an important key to know whether or not each item has met the process stages of thinking and competencies in the 21st century. Realizing the significance of English school nationally standardized examination, it is important to conduct the evaluation especially the applied higher order thinking skills on the test items as the main topic of this study entitled "The Evaluation of Higher Order Thinking Skills in English School Nationally Standardized Examination at State Senior High School 6 Semarang".

1.2 Reasons for Choosing the Topic

Some reasons for choosing these topics are first, in curriculum 2013 assessment of higher order thinking skills is one of the demands of skills in 21st-century learning, namely critical thinking, creativity, communication, and collaboration. The students will not be competitive if they are not trained in 21st-century life skills such as making comparisons, making data analysis,

making conclusions, solving problems and applying their knowledge to real-life contexts. There are three groups of competencies needed in 21st-century skills, namely: having good characters, having numbers of competencies, and mastering literacy.

Second, assessment of higher order thinking skills encourages students to do critical thinking and creative thinking so that they are do not stare on one pattern of answers from the memorization process but mastering scientific concepts. Then student`s misperception about higher order thinking skills should be corrected. The teacher`s and student`s perception of higher order thinking skills test items similar with difficult test items, actually higher order thinking skills test items are different with the difficulty level of items, it is about the thinking process to get the solution of the items to find out the correct answer. Higher order thinking skills are one component of creative thinking skills and critical thinking.

Third, in school nationally standardized examination, the proportion of higher order thinking skills test items is usually not matched with the blueprint of graduate competence standard (SKL). The teacher is difficult to construct the higher order thinking skills test items and choose the material that appropriates for the student to develop their thinking skills. Finally, there are some interferences of the teacher`s forum and school to apply higher order thinking skills such as there is a rule from nationally standardized education board that of the proportion of the higher order thinking skills around 10-15% enclose in the school nationally standardized examination, it

means that not all skills and the materials be covered as higher order thinking skills.

In this case, the teacher or teachers' forum should insert higher order thinking skills test items in school nationally standardized examination that can help students to achieve critical thinking. Therefore, school nationally standardized examination need to analyze and evaluate that higher order thinking skills have been applied so that the objectives of the assessment program can be achieved.

1.3 Research Problems

In order to clarify the problems that would be investigated, the research questions are formulated as follows:

1) How is the realization of higher order thinking skills in the listening questions of English school nationally standardized examination at state senior high school 6 Semarang?
2) How is the realization of higher order thinking skills in the speaking questions of English school nationally standardized examination at state senior high school 6 Semarang?
3) How is the realization of higher order thinking skills in the reading questions of English school nationally standardized examination at state senior high school 6 Semarang?
4) How is the realization of higher order thinking skills in the writing questions of English school nationally standardized examination at state senior high school 6 Semarang?

5) How is the content validity of English school nationally standardized examination at state senior high school 6 Semarang?
6) How is the reliability of English school nationally standardized examination at state senior high school 6 Semarang?
7) How is the practicality of English school nationally standardized examination at state senior high school 6 Semarang?

1.4 Objectives of the Study

1) to analyze the English school nationally standardized examination in order to explain higher order thinking skills of the listening questions at state senior high school 6 Semarang;
2) to analyze the English school nationally standardized examination in order to explain higher order thinking skills of the speaking questions at state senior high school 6 Semarang;
3) to analyze the English school nationally standardized examination in order to explain higher order thinking skills of reading questions at state senior high school 6 Semarang;
4) to analyze the English school nationally standardized examination in order to explain higher order thinking skills of the writing questions at state senior high school 6 Semarang;
5) to analyze the English school nationally standardized examination in order to explain the content validity of it;
6) to analyze the English school nationally standardized examination in order to explain its reliability;

7) to analyze the English school nationally standardized examination in order to explain its practicality;

1.5 Significance of the Study

The higher order thinking skills of the listening questions are explained so that theoretically, higher order thinking skills of listening can help the students to talk with others more clearly and with understanding. Practically, higher order thinking skills can help the students become better communicators and questioners. Pedagogically using the skills provide much of the information and data learners receive in language learning.

The higher order thinking skills of the speaking questions are explained so that theoretically, students can communicate both spoken and written. Practically student be able to communicate clearly and collaboration with the others and students be able to communicate their ideas effectively using various media and techniques. And pedagogically by having skilled speaking ability students are able to think creatively, work creatively and create innovation.

The higher order thinking skills of the reading questions are explained so that theoretically reading will improve the ability to understand and improve thinking skills, increase creativity and get acquainted with new ideas. Practically student be able to communicate clearly and collaboration with the others. Pedagogically reading literacy can encourage higher order thinking skills, enhance creativity, and build students' independence to solve problems.

The higher order thinking skills of the writing questions are explained so that theoretically student be able to know and mastery the system of grammar rules, mastery of linguistic aspects, mastery of discourse includes abilities arrange or organize ideas in a form of speech cohesion and coherence, and mastery of strategies in the form of abilities using verbal and nonverbal strategies to overcome various kinds of gaps that occur between the speaker/writer and the listener or reader. Practically students are able to access information effectively (information sources) and efficiently (time) evaluate information that will be used and manage information accurately and effectively to overcome problems.

The content validity of English school nationally standardized examination is explained so that theoretically, ensures whether the content is sufficiently representative and comprehensive for the test to be a valid measure of what it is supposed to measure. Practically, content validity reflects the area or the overall concept being measured.

The reliability of English school nationally standardized examination is explained so that theoretically provides information about the stability, repeatability and consistently reports the same outcomes. Practically, it is beneficial to have more items, and the types of questions to increase the test score reliability.

The practicality of English school nationally standardized examination is explained so that theoretically provides information about appropriate time constraints, easy to administer and specific and time-efficient in the

scoring/evaluation. Practically, it is prohibitively expensive and impractical in conducting the assessment program.

1.6 Scope of the Study

The researcher limits this study to make the problem not be wider. The scope of this study is the analysis of higher order thinking skills test items in the school nationally standardized examination. Then, this research will evaluate higher-order thinking skills in English school nationally standardized examination at state senior high school 6 Semarang in academic year 2018/2019.

1.7 Definitions of Key Terminologies

This section contains key terms related to this research to avoid misunderstanding among the readers. The key terms involve curriculum 2013, graduate competency standard, the assessment standard, school nationally standardized examination, higher order thinking skills, evaluation, and state senior high school 6 Semarang. Below is a brief definition of each key term.

1) Curriculum 2013

Nation & Macalister (2010) define curriculum as guidance in designing courses that consist of outer circle namely principles, environment, and needs that involve practical and theoretical considerations that will have a major effect in guiding the actual process of course production. In addition (Nunan, 1993:8) states curriculum is concerned with planning, implementation, evaluation, management, and administration of education

programs. ‘Syllabus’, on the other hand, focuses more narrowly on the selection and grading of content. Prihantoro (2015) argues that the curriculum is a document that plans the quality of learning outcomes and the educational processes that must be owned and experienced by the learners. According to the Oxford advanced learners dictionary curriculum is the subjects comprising a course of study in a school or college: course components of the school curriculum. Based on the definitions above it can be concluded that curriculum is a guideline to conduct the learning activities in order to achieve the objectives through some component and principles.

2) Graduate competency standard

Prihantoro (2015) argues that the learners are expected to improve and to balance between the soft skills and hard skills that include aspects of competencies of attitudes, skills, and knowledge. Mulyasa (2005, p.91) argues that Graduates' Competency Standards (SKL) are used as guidelines for assessment in determining the graduation of students from educational units. Graduate competency standard include the competencies of all subjects or groups of subjects. In addition the regulation of the Minister of Education and Culture number 20, 2016 stated that graduate competency standard is criteria regarding the qualifications of graduates' abilities which include attitudes, knowledge, and skills. The graduate competency standard consists of qualifications criteria of students' abilities that are expected to be achieved after

completing their study in the education unit. So, graduate competency standard is minimum qualifications criteria include attitudes, skills, and knowledge that should be achieved by the learners after complete their study from the educational unit.

3) The assessment standard

According to Fenton (1996) assessment is the collection of relevant information that may be relied on for making decisions. In addition Bachman (1990, p.18) argues that assessment is often considered as an important instructional step. On the other hand the regulation of the Minister of Education and Culture number 23, 2016 stated that the educational assessment standard is a criterion regarding the scope, objectives, benefits, principles, mechanisms, procedures, and instruments for evaluating student learning outcomes that are used as a basis in evaluating student learning outcomes in primary and secondary education. The Oxford advanced learners dictionary defines assessment is the action of assessing someone or something. From the definition above, it can be concluded that the standard of assessment is a process includes procedures used to gather information, interpret facts and make judgments in making policies.

4) School nationally standardized examination

The regulation of the Minister of Education and Culture number 4, 2018 stated that School Nationally Standardized Examination is an activity to measure student competency achievement carried out by the Education

Unit refers to Graduates' Competency Standards to gain recognition for learning achievement. The nationally standardized examination is the examination of junior and senior high school level that 75% test items constructed by the teacher forum and 25% anchor items were made by nationally education standards board.

5) Higher Order Thinking Skills

Brookhart (2010, p.5) states that higher-order thinking conceived of as the top end of Bloom's cognitive taxonomy. In addition, Anderson & Krathwohl (2001, p.21) define higher-order thinking skills is specified in three-dimensional thinking consist of analyzing, evaluating and creating. According to Heong, et al. (2011) higher-order thinking is using the thinking widely to find a new challenge. Higher order thinking demands someone apply new information or knowledge that he has got and manipulates the information to reach the possibility of an answer in a new situation. Oxford advanced learner`s dictionary defines higher order thinking skill as involving reasoning of a high level of thinking. From the definitions above, it can be inferred that higher order thinking skill is the ability to answer or solving a problem through the thinking process.

6) The Evaluation

Brown and Rodgers (2002, p. 289) define evaluation as the process of seeking to establish the value of something for some purpose. To achieve this, evaluative processes on different fields of curriculum ranging from learning, teaching and assessing should be carried out to find out the

strengths and weaknesses as well. Stufflebeam& Shinkfield (2007, p.698) stated that evaluation is the systematic process of delineating, obtaining, reporting, and applying descriptive and judgmental information about some object`s merit, worth, probity (moral correctness), feasibility, safety, significance, or equity. According to Sadler (2012), evaluation covered appraisals (a neutral term) of student learning as well as of curriculum reforms and educational projects and programs. According to Oxford advanced learners dictionary evaluation is a process of collecting data from various observations, backgrounds, and training of evaluators in order to achieve a goal. In other words, evaluation is an action or a process to determine the value of something.

7) State Senior High School 6 Semarang

State senior high school 6 Semarang located in Jl. Ronggolawe No 4 Semarang, Central Java. It is one of the schools that have a big name, therefore state senior high school 6 Semarang became one of the favorite schools especially for the people of Semarang. State senior high school 6 Semarang is well known with various achievements both academic and non-academic. As a school that already has a name, it is not surprising that state senior high school 6 Semarang is often appointed to represent Semarang in various academic and non-academic events.

1.8 The Outline of the Thesis

This research report consists of five chapters. They are chapter one to chapter five. Chapter one is the introduction, chapter two is a review of related

literature, chapter three is research methodology, chapter four is findings and discussion and chapter five is conclusions and suggestion.

In chapter one, some basic elements of the study are presented. They are; background of the study, reasons for choosing the topic, research questions, objectives of the study, significance of the study, definitions of the key terminologies, and outline of the report. In curriculum 2013, the government develops the quality of education through enhancing student`s potential for competitiveness in the 21st century. It is in line with the graduate competency standard that focuses on three aspects, they are attitude, knowledge, and skills. Especially in the knowledge domain, the learners expected to have factual, conceptual, procedural, and metacognitive knowledge at the technical, specific, detailed, and complex level related to science, technology, art, culture, and humanities. the learners able to connect the knowledge above in the context of oneself, family, school, society, and the surrounding natural environment, nation, state, and regional and international. The 21st-century skills that should be mastery by the students are literacy, 4Cs, and characters. This research focuses 4Cs that consist of critical thinking, creative, communication, and collaboration. In this case, higher order thinking skills as a part of critical thinking. To assess higher order thinking skills, questions are designed that students answer questions through thinking processes in line with operational verbs in Bloom's taxonomy, both in terms of knowledge, attitudes, and skills. Assessment of higher order thinking skills in curriculum 2013 is needed as an effort to

improve the quality of learning and improve the quality of graduates. Theoretically, this study will enlarge a theoretical review on the study higher order thinking skills in the assessment program. Practically, this research will be a model of implementation higher order thinking skills in the assessment program. Higher order thinking skills should be accustomed to the students in the learning process and also in the assessment program in order to achieve the graduate competency standard. This research also intended to help the English teacher in constructing, choose the material and applied good higher order thinking skills in the assessment program.

Chapter two review of related literature deal with reviews of previous studies, reviews of theoretical studies, and theoretical framework. In the part of reviews of previous studies, the writer took the sixties article journal from the previous researcher as references in undertaking this research. It also contains the reviews of theoretical studies which elaborate some theories related to the topic which are going to be used in this research. They relate to implementation higher order thinking skills in assessment, the development of higher order thinking skills to encourage student`s creativity, higher order thinking skills in education and pedagogical process, the evaluation of higher order thinking skills in various subjects, developing instrument for assessing higher order thinking skills, higher order thinking skills in national examination, higher order thinking skills and learning process, learning style and strategies to enhance higher order thinking skills, higher order thinking skills based on Bloom`s and Marzano`s Taxonomy, fostering students' higher

order thinking skills, and integrating higher order thinking skills in language classroom. Three theories from experts (Bloom, Barrett, and Marzano) are compared in order to be used as the basic theory in this research. As a result, Bloom`s theories of higher order thinking skill is chosen because it is the most relevant and more comprehensible. The last is the theoretical framework. It explains the scheme of how this study is conducted.

Chapter three relates to the research methodology. It consists of research assumptions, the subject of the study, the object of the study, roles of the researcher, instruments of the study, methods of collecting data, methods of analyzing data, and triangulation. The writer assumes that there will be some variation in English school nationally standardized examination test items with a different level of higher order thinking skills. The content validity and reliability are high and the practicality fulfills the requirements. this research is mix method because the research uses descriptive qualitative and also statistical formulation to answer the research problems. The subject of this study is English school nationally standardized examination while the object of the study is higher order thinking skills of it. The data of this study is taken from the documentation of the test item paper sheet and students answer sheet. Further, the researcher uses analysis card and statistical formulation of KR-21 and alpha Cronbach as the instruments and will be completed with the interview. Building trustworthiness in this study was conducted through triangulation in the form of the expert’s judgment.

Chapter four is findings and discussion. Findings part of this study explains the result of the study while the discussion part explains how the findings of this study relate to the theory used and previous studies. From this study, it is known that the realization of higher order thinking skills in English school nationally standardized examination is found. The higher order thinking conveys the listening, reading, and writing. It also found that the content validity is high, the reliability is unreliable and the practicality fulfills the requirements. From the data analysis, it is found the proportion of higher-order thinking in English school nationally standardized examination fulfills the requirements. There is also some variation of higher order thinking skills in the listening, reading and speaking questions but it is not found in the speaking questions even though the speaking skills tested simultaneously with listening in the form spoken written in English school nationally standardized examination. The content validity of the English school nationally standardized examination is high because of the representativeness of the basic competency in the test item is high. The reliability is unreliable and the practicality fulfills the requirements as the practical assessment program. In the discussion part, the writer discusses the results of the study and relate them to the theory used in the study. Further, the writer also tries to relate the results of the study with some of the previous research.

The last chapter, chapter five is the conclusion. It presents the conclusion and suggestions for the research. In conclusion part, the researcher concludes that the realization of higher order thinking skills convey listening,

reading and writing with the various levels of analyzing, evaluating and creating. The evaluation of the higher order thinking skills in English school nationally standardized examination also has high content validity but reliability is unreliable and also fulfills the requirement as the practical assessment program. Meanwhile, in the suggestion part, the researcher suggests that the higher order thinking skills proportion should be improved so that the learners accustomed to working the higher order thinking skills.

CHAPTER II

REVIEW OF RELATED LITERATURE

This chapter includes some reviews of previous studies, reviews of theoretical studies and theoretical framework.

2.1 Reviews of Previous Studies

There are several studies that has conducted related to this study. The function of the previous studies is to support the making process of this study. The previous studies grouping based on a similar topic into eleven groups, they are the implementation curriculum 2013, graduates` competencies standard, implementation higher order thinking skills in assessment, the school nationally standardized examination, higher order thinking skills in education, the research and development of higher order thinking skills, the various strategies of implementation higher order thinking skills, higher order thinking skills in various subject, higher order thinking skills in English subject and higher order thinking skills in national examination. Each group of the previous study explains such as follow:

The research about the evaluation of the implementation curriculum 2013 has conducted by Yulianto et al (2014); Qomariyah (2014); and Rumahlatua (2016); Jaedun et al (2014); Mayangsari & Santoso (2016); Nababan et al (2017); and Thoyyibah, Hartono & Bharati (2019). These studies are basically about the evaluation of the implementation of curriculum 2013. The third first studies about the evaluation of the

implementation of curriculum 2013 based on the readiness of the teachers. The teachers' readiness in implementing curriculum 2013 is in less prepared conditions. This is shown by the teachers' inadequate understanding of the principles, procedures, and techniques of the an authentic assessment, and the given tasks for the students do not portray the authentic task. The teacher also difficult to integrate between the use of ICT media, the use of scientific approaches in the learning process and conducting authentic assessment. Further Nababan et al (2017) evaluate curriculum 2013 based on the suitability of curriculum document with its implementation in SMA Negeri 7 Binjai, that there are some gaps in each stage of the implementation of the curriculum 2013. Furthermore, Mayangsari & Santoso (2016) focuses on the implementation curriculum 2013 in Kesatrian I senior high school Semarang has been implemented well but not optimum yet. Last but not least Thoyyibah, Hartono & Bharati (2019) study about the implementation of character education in English teaching-learning in curriculum 2013 that character education has implemented in some character values. The similarity of these research is curriculum 2013 as the main topic of the study while in this research curriculum 2013 as the big umbrella which is higher order thinking skills is a part of curriculum 2013.

The studies of graduates` competencies standard has conducted by Widiyanto (2013); Hidayati (2015); Fitri (2018); Asviri et al (2017); Ghina et al, (2017); Yanto et al (2018); Rosanita (2016); Pawero (2017);

Wulansari et al (2017); Maesaroh (2018); Junanto & Fajrin (2018). These researchers examine the graduate competencies standard in some various fields they are competency gaps between graduate competency standard and the needs of the industrialized, this study has also shown the effort and strategy increasing the quality of graduates in completing graduate competency standard. Content standard and graduate competency standard in religion subject, the evaluation of graduate competency standard in the PG-PAUD study program, and improving graduate competency in madrasah. Widiyanto, (2013) focuses on graduates` competency of vocational high school majority in business and management, the competence of SMK graduates needs for the workplace are knowledge, skill, attitude and other. Hidayati (2015) discuss the relevance of the graduate competency standard and the needs of the industrialized. Meanwhile Ghina et al, (2017) evaluating the level of entrepreneurship education's effectiveness through the application of a more systematic framework and relating this to a set of competencies expected on the part of graduates. The last study conducted by Yanto et al (2018) develops international competency indicators of accounting graduates. Graduates' competency standards and content standards appear to be very complex. Both are substantive and procedural. Then this also affects the technical implementation, so the quality of education must be improved in Indonesia by making clear and definite goals without ignoring the diverse potential, geographical, demographic and socio-economic conditions of Indonesia.

On the other hand in PG-PAUD study program, the four competencies of graduates namely professional, personal, social, and pedagogical competencies are in a good category however, there is still a need for improvement and improvement on professional and pedagogic competence. The competency of graduates in Madrasah can be seen from the results of national examinations over three years that increase. The policy to improve Madrasah graduate competency is done by increasing the potential of teachers and students and referring to the graduate competency standards. The similarity of this research is the graduate competency standard as the main topic while in this research graduate competency standard as the basis of the implementation of higher order thinking skills in this research.

The current assessment system requires assessments based on real situations in everyday life, students are expected to be able to apply the concepts of classroom learning to solve problems. The contextual problems faced by the world community recently are learned to relate, interpret, apply and integrate knowledge in classroom learning to solve problems in a real context. Studies of implementation higher order thinking a skill in the assessment has conducted by Widihastuti (2013); Julianingsih, Rosidin & Wahyudi (2017); Abosalem, (2016); Widana, (2017); Setiawan & Bharati, (2018); Toyoda, (2018) and Widana et al, (2018). These studies provide knowledge and understanding to the teachers about the concept and characteristics of the higher order thinking

skills in the assessment program. Generally, the assessment methods used in schools ask the students to recall information or to do routine question, which will not help students in improving their higher-order thinking skills. The impact of the integration of portfolio assessment, project-based speaking assessment, assessment HOTs through critical thinking, and assessment intercultural learning. These various assessments could improve students HOTS. The similarities of this research with these previous researches are; the first similarity is conducting the assessment of higher order thinking skills. Second, it uses Bloom`s taxonomy as the basis for determining higher order thinking skills. The differences of these study with this research are, first these studies about implementation assessment of higher order thinking skills while this research is about the evaluation of higher order thinking skills assessment. Besides, these studies focus on mathematics, accountancy and computer, intercultural learning, authentic and portfolio assessment, this research focuses on English summative assessment of the school nationally standardized examination.

The discussion of the nationally standardized examination has conducted by Herkusumo, (2011); Suwandi, (2013); Mulyadi & Suprayekti (2013); Otaya (2014); and Muzayanah, (2015). In order, these studies examine about the value of the subjects to be tested nationally among the province, among packages at UASBN also examine the actual ability of the participants UASBN based on the conversion of the values that have been synchronized. The correlation PAIKEM instructional

strategies and nationally standard school examination score. Then the evaluation of conducting school nationally standardized examination. And the rest studies are about the religious school nationally standardized examination. Overall the preparation of UASBN organizers; the level of achievement of socialization; implementation of training and preparation of UASBN questions; duplicating and distribution of UASBN questions; and supervision system are very well conducted. The studies about the religious school nationally standardized examination focus on the quality items at USBN. The similarities the previous study with the study that the writer will conduct is about school national standardized examination, but this study will focus on the English subject.

Further study of higher order thinking skills in education has conducted by Subadar (2017); Rapih & Sutaryadi (2018); Sumaryanta (2018); Fanani (2018) and Sofyan (2019). These studies focus on the higher order thinking skills and its relation with the curriculum 2013 involve character education, teacher's perspective, assessment and strategy of developing higher order thinking skill`s test. Strengthening character education based on higher order thinking skills can be achieved by training the students through contextual and meaningful classroom learning. Generally, the concepts of higher order thinking skills are understood well by the elementary school teacher. They also believe that higher order thinking skill can be implemented in elementary school, unfortunately, the teacher faced the difficulties in delivery the learning material and

designing instructional media. In conducting the assessment program especially for mathematics subject, HOTS assessment can be done for various mathematics subject with various difficulty levels. The HOTS assessment in Mathematics can be done by contextual or non-contextual questions in the form of an essay or multiple choice. While the strategy for preparing HOTS questions is done by involving all stakeholders in the education sector from the central level to the regions, and education units, in accordance with their respective main tasks and authorities. The teacher must have the knowledge and expertise to support their work so that they can develop high order thinking skills of the students. The assessment developed by the teacher is expected to be able to encourage the improvement of high-level thinking skills, enhance creativity, and build students' independence to complete the problems. The similarity of these studies and the present study is curriculum 2013 as a big umbrella while the difference between these studies and the present study is the focus of the study. The previous studies focus on HOTS and its relation with curriculum 2013 while the present study focuses on the assessment of higher order thinking skills.

Research and development of higher order thinking skills in preparing learning material, module, and student`s worksheet to increase the learning outcome and student`s characters has conducted by Widodo & Sri (2013); Prastiwi et al (2016); Kristianingsih, Wijayati & Sudarmin (2016); Anisah & Lastuti (2018); Pratiwi et al (2017); Nuraini et al (2018)

and Fanani & Dian (2018). The research of the application of higher-order thinking skills based on problem-based instruction can increase student activity and the character of students who ultimately also improve student learning outcomes. The application of HOT-PBI can improve the interaction of students and student teachers. Students are encouraged to ask the teacher, propose ideas and the formation of courage to face difficult questions. Meanwhile, the research and development of the learning material, module and student's worksheet oriented with HOTS were developed has a very good as a learning device and worthy to be used as learning materials in composing HOTS assessment instrument. The quality of the results of the development categorized as good. The similarity of these studies and the writer's research is about higher order thinking skills. These studies about research and development module and worksheet on the basis higher order thinking skills to increase the learning outcome and student`s characters while in this research focuses on the evaluation higher order thinking skills in the assessment program.

Research of higher order thinking skills and the learning strategies has conducted by Wardhani (2018); Slameto (2018); Syafa`ah & Handayani (2015); Riadi (2016); Handayani & Priatmoko (2012); Mustapa (2014) and Wahid & Karimah (2018). These studies investigate the various strategies to increase students' higher order thinking skills. The strategies are thinking actively, problem-solving, problem-based learning, cooperative learning, and their influence to increase the students HOTS.

Those learning strategies have a positive relationship and also an alternative method to increase students' HOTS. The higher order thinking skills that could be improved are creating, problem-solving, evaluating, analyzing, and critical thinking skills. HOTS can be integrated by not restricting students from submitting ideas or opinions through inquiry. In addition, HOTS can be integrated by asking students to look for various alternative answers. The conclusion of these researches points out that the selection of the strategies is crucial to increase student's higher order thinking skills. The similarity of this research and present study is highlighted the higher order thinking skills. On the other hand, the previous studies focus on the strategy in teaching-learning while this research focuses on the evaluation of assessment program.

Studies of higher order thinking skills in various subject has conducted by Yuniar, Rakhmat & Saepulrohman (2015); Mulyaningsih & Itaristanti (2018); Sari, Fitri & Hendriyani (2018); Ngongo, Relmasira & Hardini (2018); Abdullah, Alberta & Ardiansyah (2018); Martalyna, Isnarto & Asikin (2018); Musrikah (2018) Setyasih, Hartono & Prasetyo (2019). These studies explain teaching higher order thinking skills in various subjects, including HOTS test items in social science in elementary school, Bahasa Indonesia, Biology, Civics, HOTS in chemistry teacher`s candidate, HOTS in students mathematics literacy, HOTS in Math for elementary school, and increase science literacy through HOTS. These researches informed that in social science elementary school higher

order thinking skills is dominant and has been fulfilled the requirement as higher order thinking skills test items. It is in line with the HOTS in Mathematics that could be increased through stimulating the students by giving test item that oriented problem solving and problem posing. Furthermore learning content-oriented HOTS in Bahasa Indonesia has been implemented well but still need to be improved. Classroom action research in Civics reveals that HOTS through small group discussion can increase the students learning outcomes. Furthermore, HOTS questions to measure teacher candidates' HOTS has been successfully developed with valid criteria and can measure the chemical teacher candidates' HOTS on ionic equilibrium in solution topic. In addition, higher-order thinking skills-oriented is qualified for students' mathematical and science literacy. The similarity the previous studies with this present study is higher order thinking skills as the main point, while the difference of these previous studies and this study is on the object of the study, this study focuses on the evaluation of higher order thinking skills in test items.

The implementation of the higher order thinking skills in English subject has conducted by Destianingsih (2016); Anasy (2016); Mahfuzah, Jufri & Fitrawati (2019); Setyarini, S., Muslim, A. B., Rukmini, D., Yuliasri, I., & Mujianto, Y. (2018); and Febrina, Usman & Muslem (2019). In order, these studies focus on designing students worksheet reading and writing oriented HOTS, HOTS in reading exercise, student's ability in answering HOTS reading questions, improving children's HOTS

and English oral competence through critical thinking while storytelling and analysis reading questions through Bloom`s taxonomy. The worksheets that oriented on higher order thinking skills for reading and writing skills, the students preferred the topic related to their daily life, movie, and teenagers' life. Then it used to design the questions and activities for reading and writing skill-oriented on higher order thinking skills. The distribution of HOTS in reading exercises on Pathway to English textbook is categorized as lower order thinking skills. It contradicts with the result of analysis HOTS in reading exercise *Bahasa Inggris SMA/MA/SMK/MAK textbook grade 11th semester 1* that HOTS is dominant in reading exercises. While the ability of the students in answering higher order thinking in reading question show fair well. Furthermore, the HOTS in storytelling takes place in young English learner through open-ended questions by giving opinions, comments, and imaginations when the young learners analyzing and evaluating the story. Last but not least The similarity of these studies is about the implementation higher order thinking skills in English based on Bloom`s taxonomy but the writer focus on the evaluation of higher order thinking skills in school nationally standardized examination.

In addition, analysis of higher order thinking skills in the national examination has conducted by Laily &Wisudawati, (2015); Winarti et al, (2015); Iffa et al, (2016); Ahmad, (2016); Elyana et al (2016). These studies focus on the identifying higher order thinking skills content in

National examination but different subject, in order these research take the subject of Chemistry for senior high school, Physics in Madrasah Aliyah, Science Physics for junior high school and English for senior high school and Physics in Madrasah Aliyah. These research identify higher order thinking skills content using dimensions of thinking by Bloom's Taxonomy. The result of these research shows that in the characteristics of the type of higher order thinking skills matter that existed at the national examination of Chemistry is the stimulus, while the ability to think critically and creatively conclusions cannot be generated. The national examination of Physics in Madrasah Aliyah is merely the memorizing and the formula applications. The Physics examination is given with the level of remembering, understanding, implementing and analyzing. The level of evaluating and creating is rarely and even never used. It is in line with the nationally examination Science Physics for junior high school that the category higher order thinking skills test items on the national exam science physics is still very low. In English nationally examination items, the items categorized into the literal level, reorganization and inferential level the tests were not enriched sufficiently with the evaluation comprehension. It is obvious that there is a shortage of items questioning students' higher order thinking skills in the exam and they are not well-treated. Last but not least higher order thinking in completing Physics national examination is generally capable of measuring higher order thinking skills of students but not well implemented. Overall, it can be

concluded that higher order thinking skills in national examination need to be modified by providing them with more question items that include higher order thinking skills.

This research is different from the previous study this research consider Bloom`s taxonomy as the basic theory of determining higher order thinking skills level. This research focuses on the realization of higher order thinking skills in English school nationally standardized examination. Besides that this research also evaluates the implementation of higher order in English school nationally standardized examination. From all these reasons above, the writer conducts the study about the evaluation of high order thinking skills in English school nationally standardized examination at state senior high school 6 Semarang.

2.2 Reviews of Theoretical Studies

The reviews of theoretical studies will explain some theories related to this research, it consists of curriculum 2013, graduate competence standard, assessment standard, school nationally standardized examination, the definition of higher order thinking skills, higher order thinking skills in English skills, and evaluation.

2.2.1 The Curriculum 2013

The curriculum 2013 has been implemented in stages and limited since the academic year 2013/2014. Curriculum 2013 is a curriculum that currently applies nationally. Based on the constitution of Indonesia number 20/2003 curriculum is a set of plans and regulations regarding

the objectives, contents, learning materials and the methods used to guide the implementation of learning activities to achieve certain educational goals (The regulation of the Minister of Education and Culture number 59, 2014). The curriculum develops attitudes, knowledge, and skills that placed Indonesian culture as the basis for developing Indonesian education also capable and useful for developing the quality of Indonesian people.

The curriculum 2013 aims to prepare Indonesian people having the ability to live as individuals and citizens who are faithful, productive, creative, innovative, and effective and able to contribute to the community, nation, and state of the world. To manifest it, in the Minister of Education and Culture regulations about standards process it was stated that the learning process in educational units was held interactively, inspiring, fun, challenging, motivating students to actively participate, and providing chance for initiatives, creativity, and independence in accordance with the talents, interests, physical and psychological development of students. Learning is directed to encourage students to find out from various sources of observation, being able to formulate problems not only solve problems. Besides that learning is directed to train students to think analytically in decision making rather than mechanistic and able to work together and collaborate in solving problems

Related to the international issue of the development of education, the curriculum 2013 was designed with various enhancements, such as on standards of content by reducing irrelevant material and deepening and expanding relevant material for students and enriched based on students' needs to think critically and analytically by following international standards. The standards of assessment improved by gradually adapting international standard assessment models. Learning outcomes assessment is expected to help students to improve higher order thinking skills (HOTS) because higher order thinking skills can encourage students to think broadly and deeply about the subject matter.

The government expects students to achieve various competencies by applying higher order thinking skills. These competencies are critical thinking, creative and innovative, communication skills, and collaboration inherent of the evaluation system in national examinations and they are called 21st-century skills. Higher order thinking skills (HOTS) are implemented because of the low ranking of the Program for Internationally Student Assessment (PISA) and Trends in the International Mathematics and Science Study (TIMSS) so that the standard nationally examination test items are tried to be improved.

The development of learning that oriented to higher order thinking skills (HOTS) is a program developed by the Ministry of

Education and Culture as an effort to improve the quality of learning and improve the quality of graduates. This program was developed as the policy of the Ministry of Education and Culture which in 2018 had integrated Character Education and learning-oriented towards higher order thinking skills.

2.2.2 The Graduate Competency Standard

Graduate competency standard is criteria regarding the qualifications of graduates' abilities which include attitudes, knowledge, and skills. The graduate competency standard consists of qualifications criteria of students' abilities that are expected to be achieved after completing their study in the education unit (The regulation of the Minister of Education and Culture number 20, 2016). To find out the achievement and appropriateness between graduate competency standard and graduates of each education unit and curriculum used in certain education units, it is necessary to carry out regular and continuous monitoring and evaluation in each period. The results obtained from monitoring and evaluations are used as input for future improvement of graduate competency standard.

In curriculum 2013 the graduate competency standards consist of three areas, they are attitudes, knowledge, and skills. Spiritual and social attitudes are the domain of competence should be having by graduates in accordance with their religion. Social attitudes are

competencies that graduate is expected to have associated with humans relations.

Graduate competency standard of curriculum 2013 in the knowledge domain includes four types of knowledge namely factual, conceptual, procedural, and metacognitive knowledge. Factual knowledge relates to facts around students. Conceptual knowledge is related to the theories about facts that have previously been studied. Procedural knowledge is knowledge about how to do things. Metacognitive knowledge is related to self-understanding, which is about one's strengths and weaknesses of the knowledge. Metacognitive knowledge is the knowledge of oneself to do something. Metacognitive knowledge is closely related to concrete skills.

Skill is the ability to do something and making something. The types of skills that students need to achieve to meet graduate competency standards are concrete skills and abstract skills. Concrete skills are skills related to direct body movements, for example imitating someone's actions, producing products by following general instructions. In this category, students are guided through instructions to perform real action. The real example of concrete skills e.g, students practice role-play in front of the class from the dialogue given by the teacher, a student read the passage after the teacher giving instruction student x read the passage loudly, please. These examples happen in the teaching and learning process. Abstract skills are skills related to

abstract actions that can be performed by students. Abstract skills refer to communicating, experimenting, associating, and creating in the knowledge domain. For example in communicating the students be able to present the results of the study from observing to reasoning in the form of written, graphics, multi-media electronics, etc. the other example is the ability to generate ideas or make new decisions.

2.2.3 The Assessment Standard

Assessment is the process of gathering and processing information to measure student learning outcomes. The educational assessment consists of the assessment of learning outcomes by educators, assessment of learning outcomes by education units and assessment of learning outcomes by the government. The educational assessment standard is a criterion regarding the scope, objectives, benefits, principles, mechanisms, procedures, and instruments for evaluating student learning outcomes that are used as a basis in evaluating student learning outcomes in primary and secondary education (The regulation of the Minister of Education and Culture number 23, 2016). The aims of assessment of learning outcomes are monitoring and evaluating the process, learning progress, and continuous improvement of student learning outcomes.

Assessment is carried out through three approaches, namely: assessment of learning, assessment for learning and assessment as learning. Assessment of learning is an assessment carried out after the

learning process is complete. This assessment is intended to determine the achievement of learning outcomes after students have finished the learning process. Various forms of summative assessment at the end of semester tests, school examination and nationally examinations are examples of assessment of learning.

Assessment for learning is carried out during the learning process and used as a basis for improving the learning process. Assessment for learning teachers provide feedback on students learning process, monitor progress, and determine learning progress. Assessment for learning is a process of assessment used by teachers to improve their performance in facilitating students. Various forms of formative assessment, for example, assignments, presentations, quizzes, are some examples of assessment for learning.

Assessment as learning is similar to assessment for learners because it is also carried out during the learning process. The difference is that assessment as learning involves students actively in the assessment; students are allowed to assess themselves or give an honest assessment of their friends. Self-assessment and peer-assessment are examples of assessment as learning. In assessment as learning students are involved in formulating criteria and rubric/guideline assessment procedures so that they know exactly what should be done to obtain maximum learning outcomes.

The assessment in curriculum 2013 consists of the assessment of attitudes (affective), knowledge (cognitive) and skills (psychomotor). Assessment of attitudes is an assessment of students' behavior during the learning process, in the classroom, and outside the classroom to develop the attitudes, behavior, and character of the student. Assessment of attitudes consists of spiritual attitudes and social. Spiritual attitude assessment is carried out form students' attitudes so that they are able to appreciate, inspired, and practice of their religion. Social attitude assessment is done to implement the student sense of respect, honest, disciplined, responsible, caring, polite, and confident in interacting effectively with the social environment and the natural environment.

Knowledge assessment is an assessment conducted to find out students' mastery that consists of factual, conceptual, and procedural knowledge and levels of thinking skills. Knowledge assessment is carried out with various assessment techniques. besides to find out mastery learning, knowledge assessment also identify the strengths and weaknesses of students' mastering knowledge in the learning process (diagnostic). So, giving feedback to students and teachers is very important, in order that the results of the assessment can be immediately used to improve the quality of learning. Knowledge assessment conducted in the form of a written test, spoken test, and assignments.

A written test is a test with questions and answers presented in writing to measure or obtain information about the students' abilities. Written tests require responses from students that can be used as a representation of their abilities. Written test instruments can be multiple choice questions, filling; short answers, true-false, matchmaking, etc. A spoken test is giving a question or question that requires students to answer orally, and it can be given classically during the learning process. Student answers can be in the form of words, phrases, sentences or paragraphs. Spoken tests foster the attitude of students to argue. The assignment is giving a task to the students to measure and/or increase students` knowledge. Assignments used to measure knowledge can be done after the learning process, while assignments used to increase knowledge are given before and/or during the learning process. Assignments can be done individually or in groups according to task characteristics.

Assessment of skills is an assessment carried out to measure the ability of students to apply knowledge in performing certain tasks in various contexts according to indicators of achieving competence. Assessment of skills can be done with a variety of techniques, including practice, product, project, and portfolio assessment. The technique used in the assessment of skills was chosen according to the characteristics of basic competence in core competencies or KI-4.

Assessment of practice demands a response of skills through activity in accordance with the demands of competence. The aspect assessed in practical assessment is the quality of the process of doing a task. Assessment of practice aims to assess students' abilities in demonstrating their skills through an activity. The practical assessment is more authentic than paper and pencil assessments because the forms of their duties reflect the abilities needed in the practice of everyday life.

Assessment of product is an assessment of students' skills in applying their knowledge into the form of a product according to the criteria that have been set both in terms of processes and final results. Product assessment is done on the quality of a product. Assessment of product aims to (1) assess students' skills in making certain products according to achieving classroom learning goals; (2) assessing mastery of skills as a condition for learning the next skill; and (3) assess students' ability to explore and develop ideas in designing and showing innovation and creation.

Assessment of the project is an activity to determine the ability of students to apply their knowledge through the completion of a project instrument in a certain period/time. Assessment of the project can be done to measure one or several basic competencies (KD) in one or several subjects. The instrument is in the form of a series of activities ranging from planning, collecting data, organizing data, processing and

presenting data, and reporting. Assessment of a portfolio is a technique for assessing aspects of skills. The portfolio is a collection of samples of the best works of basic competencies (KD) on KI-4. The sample is collected from products that are produced from the technique project and product assessment.

2.2.4 School Nationally Standardized Examination

The assessments of learning outcomes are conducted to diagnose the strengths and weaknesses of students learning. To achieved that purposes can be used in various forms and instruments. Assessments can be carried out; spoken, written, performance or assignments such as projects. The quality of the assessment instrument is an important factor in its implementation. Therefore, developing instruments need to be continuously improved so that information obtained from the results can be responsible. Nationally Examination (UN) and School Nationally Standardized Examination (USBN) cannot be separated from the national education system. The National Examination is an evaluation system for the standards of primary and secondary national education and the equal quality of education levels between regions carried out by the Education Assessment Center.

School nationally standardized examination (USBN) is the education unit nationally standardized summative test. Because the results of USBN determine the graduation from the education unit, the USBN test items are expected to fulfill good instrument requirements to

provide valid and objective information. The invalid and unreliable test items, giving information that is not appropriate with student achievements so that, it can harm students and provide inappropriate or misleading information for decision-makers. Since academic year 2016/2017, school examinations in several subjects enhanced become school nationally standardized examination (USBN) at the level of junior high school and senior high school. The preparation of USBN test items based on the blueprint set by the board of education national standard (BSNP). In some subjects, 20% - 25% of USBN test items come from the Ministry of Education and Culture, and then 75% - 80% of test items are prepared by the teachers who are then consolidated in the teachers' forum (MGMP). This blueprint contains indicators, constructing test items and constructing practice tests. The form of the USBN test items is written test consists of multiple choices and essays.

In conducting the school nationally standardized examination, teachers can insert some higher-order thinking skills points. Based on Widana (2017) here are some roles high orders thinking skills issues in improving the quality of assessment.

1) Preparing the competencies of learners to meet the 21st-century

   The assessment conducted by the educational unit is expected to set up learners to have some of the competencies needed in the 21st century. There are 3 groups of competencies needed in the 21st century skills, namely: a) good character (religious, curiosity, social

and cultural sensitivity, adaptable, and competitiveness); b) competencies (critical and creative thinking, problem-solving, collaboration, and communication); and c) literacy includes thinking skills using knowledge resources in printed, visual, digital, and auditory. Higher order thinking skills test items in assessment can train learners to improve ability and skill by following the demands of 21st-century competence above. Through assessment based on higher order thinking skills questions, critical thinking skills, creativity, and self-reliance, will be built through practice activities to solve real problems in daily life (problem-solving).

2) Improve a sense of belonging and care for the region improvement

In the assessment, teachers are expected to develop higher order thinking skills issues creatively appropriate with situations and conditions in their respective areas. Teacher's creativity is important in the selection of stimulus based problem areas in the educational environment. Various problems that occur in the area can be lifted as a stimulus contextually. Thus the stimulus was chosen by the teacher in the higher order thinking skills test items becomes very interesting because it can be seen and felt directly by the learners, the higher order thinking skills test items in school examination can enhance the sense of belonging and care for local potentials that exist in the student`s environment. So that learners feel to take part to solve the various problems that arise in their region.

3) Increasing students' learning motivation

Formal education should be able to answer the challenges in society. The knowledge that is learned in the classroom, directly related to problem-solving in the community. Thus students feel that the learning material obtained in the classroom is useful as a provision to get involved in the community. Challenges that occur in the community can be used as a contextual and interesting stimulus in the assessment, so higher order thinking skills questions are expected to increase students' learning motivation.

4) Improving the quality of assessment

The quality of the assessment will improve the quality of education. By training students to answer higher order thinking skills questions, it is expected that students can think critically and creatively.

2.2.5 Higher Order Thinking skills

2.2.5.1 Bloom`s Taxonomy

Bloom's Taxonomy first introduced in 1956 when he chaired a committee to outline educational objectives and assessments. The taxonomy is separated into three functional categories: knowledge (cognitive), attitudes (affective) and skills (psychomotor). These three, higher order thinking skills are part of the first category, the knowledge (cognitive) domain, which consists of six further classifications: knowledge, comprehension, application, analysis, synthesis, and

evaluation. Each category has a subcategory to describe its function; However, the taxonomy is not a true taxonomy because the process is not always hierarchical. It means that there is no linear way, particularly in the higher levels so, learners can jump or skip levels freely.

Then, this original framework was modified by Anderson et al. (2001, p.28), the revised taxonomy include: remembering, understanding, applying, analyzing, evaluating, and creating. The revised taxonomy can be seen in figure 1:

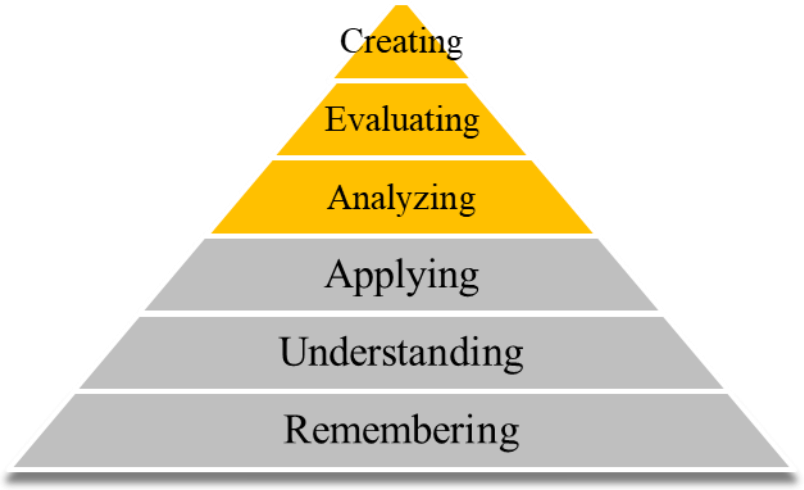

Figure 1. The Revised Bloom`s Taxonomy

The differences in both versions are clear, but not as completely as they may appear. The first level in the original taxonomy, knowledge, is become remembering to make this stage more functional. Students are not only expected to have facts but also to remember them accurately for use. There is an increase in cognitive processes for example learners are not required to know a concept only but also remember the concept they studied. Then the comprehension is changed to understanding for the

reason that most teachers prefer the term understanding instead of comprehension. Finally, the synthesis domain no longer exists, but actually, it is included in creating`s level, evaluating is positioned to where synthesis was in the previous taxonomy and synthesis is changed to creating in the last position. The reason for this change is that the process of the taxonomy is meant in these final terms to reflect a process by where evaluations are made (mentally) and items are created (production) reflecting previously relevant terms (Krathwohl, 2002, p.212).

The conforming higher order thinking skills, thinking level is seen from Bloom's revised cognitive domain they are analyzing, evaluating, and creating. To assess the thinking skills of learners, the outcome learning is designed in such a way that learners answer the questions through thinking processes that appropriate with the operational verb in Bloom's taxonomy of knowledge (cognitive), attitudes (affective) and skills (psychomotor) domains.

Several operational verbs are similar in some cognitive domains, for example, the verb explains in the knowledge level and explains in understanding level, as well as conclude in the level of analysis and concluded in the evaluation level. This difference can be seen in the form of the operational verb in figure 2:

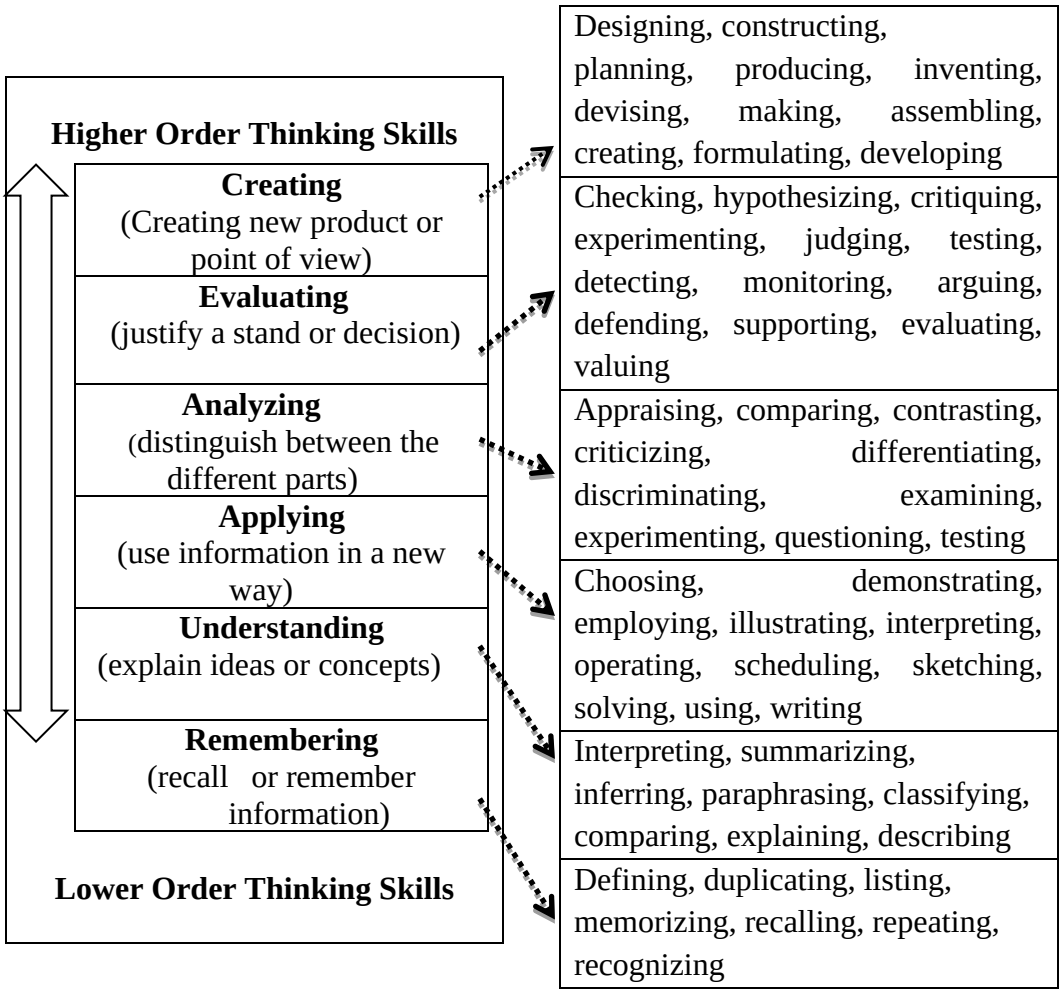

Figure 2. The Operational Verbs in Cognitive Domain

The dimension of thinking processes in Bloom's taxonomy that revised by Anderson & Krathwohl (2001, p.28), consisting the ability to remember, (C1), understanding, (C2), applying (C3), analyzing (C4), evaluating (C5), and creating (C6). Higher order thinking skills questions generally measure the ability to analyzing (C4), evaluating (C5), and creating (C6).

The selection of operational verbs to formulate indicator of higher order thinking skills, should not be stuck on grouping instructional verbs. As a word, the 'decisive' work on Bloom's Taxonomy lies in C2 and C3 level. In the context of writing higher order thinking skills questions, the 'deciding' verbs may be present in C5 (evaluating) if the decision is preceded by the process of analyzing information in the stimulus and then learners are asked to determine the best decision. Even the operational verbs 'determine' can be classified C6 (creating) when the question demands the ability to compose a new troubleshooting strategy. Thus, the dimension of the instructional verb is strongly influenced by what process of thinking is needed to answer the question.

2.2.5.2 Barret`s Taxonomy

The other taxonomy is proposed by Barret, unlike Bloom`s that focuses on the cognitive, Barret`s emphasizes on comprehension skill, Barrett’s taxonomy of comprehension, discusses the different levels of comprehension namely: literal comprehension, reorganization, inferential comprehension, evaluation, and appreciation. Barrett as cited in Vethamani (2008) assumes that learners move from the literal comprehension of understanding to another until the learner understands and appreciates the cognitive and aesthetic aspects of the material. Literal comprehension and reorganization comprehension, deal with the facts as presented orally or in the books the students have read, and thus result in closed questions that have a single correct response. Inferential

comprehension is showed when students employ the ideas and information explicitly stated in a viewing material, students' intuition and personal experiences as bases in making intelligent guesses and hypothesis. Evaluation comprehension refers to assessing the language and effect of the material in the light of appropriate criteria. It needs responses which indicate that an evaluative judgment has been made by comparing ideas. Appreciation comprehension deals with psychological and aesthetic responses. It implies emotional responses to content, plot or theme, sensitivity to various literary genres, identification with characters and incidents, reaction to the author's use of language, and response to generated images. The remaining categories always involve the student's background knowledge. Consequently, many different, but correct, responses will emerge since each student owns a different background of home, family, friends, and learning process. These categories, therefore, lead to the development of open-ended questions which require students to use higher order thinking skills. The dimensions of comprehension skills based on Barret can be seen in figure 3

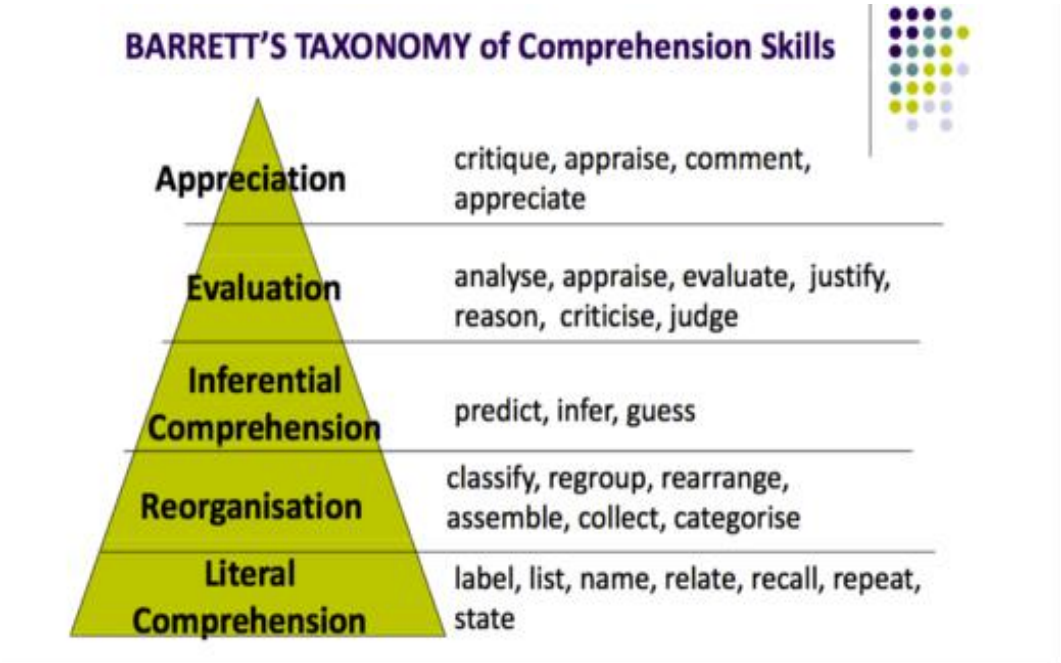

Figure 3. The Dimension of Barret`s taxonomy

The aspect of the Barrett taxonomy, according to Armbruster & Ostertag as a cited in Ahmad (2016), is the subdivision of categories based on the specific type of information targeted by the question (e.g. recognizing and recalling main ideas, inferring cause and effect relationships, identification with characters and incidents). It promotes the use of Barrett's taxonomy as a model for constructing questions on a category of levels as well as for judging questions that have already been created. It can be used to evaluate students' comprehension of the text. According to Ahmad (2016), Bloom's taxonomy of higher thinking skills sheds light on Barrett's comprehension as illustrated in table 2.1

Table 2.1. Bloom`s, Barret`s Taxonomy and two levels of thinking skills

<table>
<tr><th>Bloom`s taxonomy</th><th>Barret`s Taxonomy</th><th>Two level of thinking skills</th></tr>
<tr><td>Knowledge</td><td>Literal comprehension or recall</td><td rowspan="7">Lower order thinking skills

Higher order thinking skills</td></tr>
<tr><td>Comprehension</td><td>Reorganization</td></tr>
<tr><td>Application</td><td rowspan="3">Inference</td></tr>
<tr><td>Analysis</td></tr>
<tr><td>Synthesis</td></tr>
<tr><td>Evaluation</td><td>Evaluation</td></tr>
<tr><td></td><td>Appreciation</td></tr>
</table>

The right column shows two categories based on the required level of cognitive operation: lower order thinking skills and higher order thinking skills. The first demands the literal comprehension or recall of factual information explicitly presented in the text. The information generally involves facts, names, dates, times, locations, lexical items, and propositions. Literal comprehension and reorganization include lower order thinking skills category because questions of literal comprehension and reorganization can be answered directly and explicitly from the text. On the other hand, higher order thinking skills require more than mere recognition or recalling information. They also facilitate moving beyond a literal understanding of the text to a more knowledge-based and global understanding of textual meaning. In other words, they require readers to read beyond the lines. Thus, inferential comprehension, evaluation, and appreciation belong to higher order thinking skills because, in order to

answer these types of question, students must use both a literal understanding of the text and their knowledge of the text's topic and related issues.

2.2.5.3 Marzano`s Taxonomy

Marzano's research on thinking skills is important to students and educators in higher education institution primarily. Marzano identifies an integrated full theory of learning based on the most recent brain-based learning literature. According to Marzano and Kendall (2007, p.28), there are three systems of thought, they are the self-system, the metacognitive system, and the cognitive system. The cognitive system consists of retrieval, comprehension, analysis and knowledge utilization. Retrieval tasks ask students to access information exactly as it was originally presented. Comprehension requires interpretation of information and aims an internalizing it. An understanding at this level is demonstrated when students can articulate and/or present a non-linguistic representation of the most essential/important characteristics of the information, concept or idea. The analysis involves extending their knowledge as they discover new relationships and applications. Knowledge utilization focuses on using knowledge to address more authentic tasks. These elements of the cognitive system build up the higher order thinking skills. Each cognitive level consists of several categories with the various operational verbs can be seen in figure.3

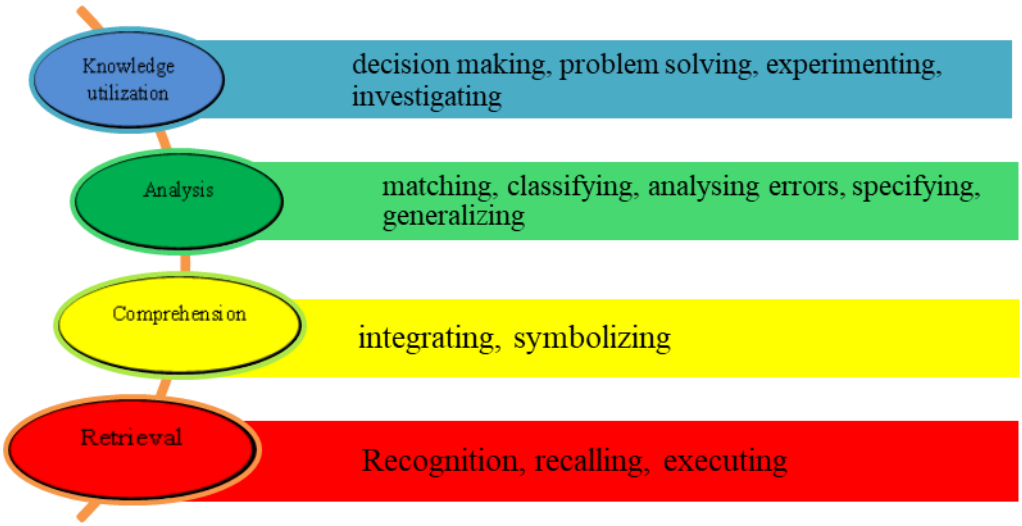

Figure 4. Dimension and Operational Verbs in Marzano Taxonomy

From the figure above there are two key features of Marzano's taxonomy make it useful in designing courses, assessments and the sequencing of instruction. First, cognitive levels are differentiated based on how much intentional cognitive effort is expended by a student as they completed the task. At one end of the spectrum, retrieval tasks typically require little to no processing as students pull information directly from working memory; on the other hand, knowledge utilization invokes the most significant intentional effort. Marzano's taxonomy is beneficial in the sense that it helps us think more seriously about the cognitive load we place on our students when asking questions/building understanding.

A second key feature of the taxonomy is taken from the hierarchical nature of the cognitive levels in Marzano's taxonomy. Not having the ability to retrieve, for example, will preclude students' ability

to successfully address comprehension tasks. A lack of comprehension will prevent the successful completion of an analysis task. The hierarchically arranged cognitive levels allow for a careful treatment of scaffolding in instruction. For example, the first exposure to new material could be dedicated to learning activities centered on retrieval or comprehension, to be expanded to higher-order thinking activities as a unit progresses.

Based on the three expert`s theories, the researcher refers to the first author Bloom`s revised taxonomy by Anderson & Krathwohl (2001) because his theory is the most relevant to this study. Thus, the researcher follows Bloom`s revised taxonomy of higher order thinking skills.

Higher order thinking skills include the ability to solve problems (problem-solving), critical thinking skills, creative thinking, the ability to argue (reasoning), and decision-making ability. A higher order thinking skill is one of the important competencies in the modern world, so it must be mastered by the students. The creativity of problem-solving in higher order thinking skills, consisting of:

a. solve unfamiliar problems;
b. evaluate the strategies used to solve the problem from a different point of views;
c. Find new models of settlement that differ from previous ways.

The difficulty level is not the same as higher order thinking. The level of difficulty in the test items is not the same with higher order

thinking skills. For example, to know the meaning of the uncommon word may have a very high degree of difficulty, but the ability to answer the problem does not include higher order thinking skills. Thus, higher order thinking skills issues are different with difficulty level. Higher order thinking skills can be trained in the classroom learning process. So that the learners have the ability to think higher order, then the learning process also provides space for learners to find concepts activity-based knowledge. Activity in the learning process will encourage learners to build creativity and critical thinking.

Higher order thinking skills issues are real-life situational assessments, where learners are expected to apply the concepts of learning in the classroom to solve problems. Contextual problems faced by the world community currently related to the global awareness, financial, economic, business and entrepreneurial literacy, civic literacy and health literacy, as well as utilization of science and technology in various aspects of life like understanding includes how the learners' skills to be related, interpret, apply and integrate science into classroom learning solve problems in a real context.

2.2.6 Higher Order Thinking in English Skills

2.2.6.1 HOTS in Listening

Listening is more than merely hearing words. Nunan (2003, p.24) states that listening is an active, purposeful process of making sense of what we hear. It means that we should interpret the meaning

of what we hear, therefore listening is an active skill. As one of the receptive skills, listening skills become an element that must be mastered first by students. Indeed, naturally, the first time humans learning the language of others through listening, so these concepts view that foreign language skill that must be learned is listening. To achieve these goals it is necessary to learn to listen through a process of high-level thinking. Realization the questions of higher order thinking skills for listening skills can be seen in the example below:

Table 2.2 The Realization of LOTS and HOTS in Listening Skill

LOTS question in listening skill	**HOTS question in listening skill**
Woman: Excuse me. Would you please come to the meeting after lunch? Man : Why not? I'm a new employee, I'm ready to go anytime. Woman: Thank you. Don't forget to bring your laptop! Man : Yes Ma'am, anything else? Woman: Okay, thank you. Man: _____	Man: Oh my god, this traffic is making me crazy. Woman: Absolutely! I think there must be an accident. Man: I think so too. It's almost seven o'clock. We'll be late for school. Woman: We'd better tell our teacher that we're coming late, because of the traffic. We hope and pray that the traffic will be fine soon and we will not miss the

Narrator: What does the man probably reply? A. Sorry, I don't understand you B. It's okay. Don't disturb me! C. You know, actually, I'm busy D. You are welcome	first lesson. Man: Okay, let's call her Narrator: What is the best solution they have? A. Make themselves crazy B. Come late to school C. Call the teacher D. Hope and pray E. Miss the first lesson
The correct choice is D. Answering this question does not need a higher order of thinking because this question is routine only needs remembering or requiring literal comprehension. Therefore, it belongs to lower order thinking skills.	The correct choice is C. Answering this question need a higher order of thinking because to answer this question it needs to infer or conclude from the dialogue. Therefore, it belongs to higher order thinking skills.

2.2.6.2 HOTS in Speaking

Curriculum 2013 emphasizes speaking as one of the important skills and must be mastered by students, especially to communicate. Communication is one of the skills needed in the 21st century.

According to Harmer (2007, p.284), speaking is the ability to speak and presupposes not only knowledge of language features, but also the ability to process information and language on the spot. In addition Kusniandang & Barathi (2016) argue that speaking skill is very important because when we learn a new language, the main goal is to have capability to use it in daily communication. Furthermore, Safari & Fitriati (2016) stated that in order to be able to speak fluently and correctly, students must master sufficient vocabulary and able to arrange their expressions so that the meaning can be conveyed correctly.

In speaking, higher order thinking skills are expected to encourage students to communicate both spoken and written. Higher order thinking skills in speaking aims to understand manage and create effective communication in various forms and contents spoken, written, and multimedia. Use the ability to express ideas, both during the discussion, inside and outside the classroom. Use spoken language that appropriates with the content and context of the conversation. In addition, oral communication requires an attitude to listen, and respect the other`s opinions, besides the knowledge of the content and context of the conversation. In the 21st century skills, communication is not limited to only one language, but it is likely to be multi-language. The manifestation of higher order thinking processes on speaking skills questions can be seen in the following example:

Table 2.3 The Realization of LOTS and HOTS in Speaking Skill

LOTS question in speaking skill	**HOTS question in speaking skill**
Describing or explaining one thing or famous person provided. Example: Describe one of the things or famous person in ten sentences communicatively. things: 1. Plane 2. Eiffel Tower 3. Candi Borobudur Famous People: 1. B.J Habibie 2. R.A Kartini	Arguing and reasoning after the students reading the narrative text about Kancil and crocodile. Example: Do you think Kancil has done the right things? and why?
Answering this question does not need a higher order of thinking because it just tells and delivers from the previous and source that students studied before. Therefore, it belongs to lower order thinking skills in understanding level.	Answering this question needs a higher order of thinking because to answer this question it needs to evaluate the Kancil`s character and giving the reason. Therefore, it belongs to higher order thinking skills in evaluating level.

2.2.6.3 HOTS in Reading

Reading is one of the basic skills of language learning. It can not be separated from other skills of language learning besides writing, speaking and listening. According to Harmer (2007, p.99), reading is useful for language acquisition. Provided that students more or less understand what they read, the more they read, the better they get at it. Based on Grellet (2004, p.7) reading is a constant process of guessing, and what one brings to the next is often more important than what one finds in it. In reading, the students should be taught to use what they know to understand unknown elements, whether these are ideas or simple words.

Furthermore, in reading the questions can be classified into the higher or the lower levels of students thinking. Lower thinking requires the students to recall, the higher level requires the students to perform in a complicated process. It means that the higher order thinking skill in the reading requires the students to use their thinking skill in a complicated process. The lower order thinking questions are more easy to be found in the test format or the exercise because they are familiar, the easiest to answer by the students. from the explanations above about the higher order thinking in reading. the examples of lower order thinking skills and higher order thinking skills questions in reading such as follow:

Table 2.4 The Realization of LOTS and HOTS in Reading Skill

LOTS question in reading skill	**HOTS question in reading skill**
Students are able to determine the use of the text provided in social life. example: - Who do you think will most likely use the text above? (the text is about a recipe) A. a cook B. a technician C. a researcher D. a teacher E. a student Where do you think you can find this kind of text?	Students are able to associate the function with another text in social life (recount text provided) example: - Which of the following has the same function as the text above? A. A report to the police about a robbery B. A story about an animal C. A letter to a principal about an activity D. A speech about a phenomenon E. A story of my childhood
Answering this question does not need a higher order of thinking because it just tells and applying the knowledge from the previous that students studied before. Therefore, it belongs to lower order thinking skills in applying level.	Answering this question needs a higher order of thinking because to answer this question it needs to associate with the other text. Therefore, it belongs to higher order thinking skills in evaluating level.

2.2.6.4 HOTS in Writing

Writing is a language skill that is used to communicate indirectly, not face to face with other people. Nunan (2003, p.88) states that thinking is the mental work of inventing ideas, thinking about how to express, and organizing them into statements and paragraphs that will be clear to a reader. Writing is a productive and expressive activity. In this writing activity, the writer must be skilled in utilizing graphology, language structure, and vocabulary. Writing skills do not come automatically, but must be learned through a lot of practice and practice regularly. Writing is a complex activity that includes integrated arms, hands, fingers and eyes. Writing is also related to understanding language and speaking skills.

Writing skills are the basis for mastering various fields of study. Listening and writing are also useful for recreation or for obtaining pleasure. Writing is the ability of a person to reduce graphic symbols that describe a language and express ideas to create a note or information on a media so that other people can hear the notes or information. The realization of higher order thinking skills in writing as follows:

Table 2.5 The Realization of LOTS and HOTS in Writing Skill

LOTS question in writing skills
Example: At present, airplanes are playing a very important role to ... (34) one place to another. People can go round the world just in a two-day flight by the world's first supersonic airliner, Concorde, which ... (35) at a height of over 180.000 meters and ... (36) a speed of over 2000 km per hour. 34. A. disjoin B. part C. connect D. separate E. Divide 35. A. goes B. flies C. jumps D. drives E. comes 36. A. cuts B. continues C. moves D. reaches E. Limits

Answering this question does not need a higher order of thinking because it just recalls from the previous material that students studied before then apply it in the text. Therefore, it belongs to lower order thinking skills in applying level.

HOTS question in writing skills
Example: Compose a letter of apology from Kancil to Crocodile!
Answering this question needs a higher order of thinking because to answer this question it needs to analyze the Kancil mistake first then evaluate that Kancil has been done the wrong thing until creating an apology letter. Therefore, it belongs to higher order thinking skills in creating`s level.

2.2.7 Evaluation

Recently evaluation has been an independent object of increased interest from the different point of view of the different discipline. In education, evaluation needs to get information about whether the teaching-learning process is being successful. Theory of evaluation is highlighted by some authors such as Stufflebeam & Shinkfield (2007), Brown and Rodgers (2002), and Government regulation number 19, 2005.

Evaluation is the framework to get the information of to what extent the objectives achieved. Stufflebeam& Shinkfield (2007, p.698)

stated that evaluation is the systematic process of delineating, obtaining, reporting, and applying descriptive and judgmental information about some object`s merit, worth, probity (moral correctness), feasibility, safety, significance, or equity. In other words, evaluation is an action or a process to determine the value of something.

The evaluation is needed to find out whether the learning objectives are achieved and the learning process that has been carried out. Brown and Rodgers (2002, p. 289) define evaluation as "the process of seeking to establish the value of something for some purpose". To achieve this, evaluative processes on different fields of curriculum ranging from learning, teaching and assessing should be carried out to find out the strengths and weaknesses as well. Moreover, Government regulation number 19 (2005) stated that Evaluation is the process of gathering and processing information to measure student`s learning outcomes.

According to Stufflebeam & Shinkfield (2007, p.21) there are three types, they are:

a) Formative evaluation ensures that a program or program activity is feasible, appropriate, and acceptable before it is fully implemented. The evaluation will uncover processes and other information for program improvement, help determine what works best and accept the fact that negative results, these provide feedback for improvement.
b) Summative evaluation determines whether program activities have been implemented as intended. Summative evaluations examine the

overall quality and outcomes of a program. These evaluations are conducted for decision-making purposes and determine whether the program has met its intended outcomes relative to its cost.

c) Formative and Summative

"Both formative and summative evaluation is needed in the development of a product or service. Too often, summative evaluation is carried out only for judging programs. This restricts development processes." (Stufflebeam & Shinkfield, 2007, p. 24).

Accordingly, Stufflebeam & Shinkfield`s definition is the most comprehensive one, it sheds the light on the process of evaluating a particular assessment taking into account the needs of the participants. Thus, the researcher follows Stufflebeam & Shinkfield theory about evaluation.

In evaluating of the assessment program, three elements that show the assessment program fulfills the characteristics as a good assessment program. The three elements are validity, reliability, and practicality.

2.2.7.1 Content Validity

Validity in social research is a very important problem because of the accuracy of the measuring instrument used. It can be interpreted that an inappropriate/inappropriate instrument will have implications for the validity of the research results themselves. The validity, according to McCowan and McCowan (1999), means the extent to which a test measures what it is supposed to measure. One way to test

the validity of the test is by assessing the test items. There are three types of validity, namely content validity, criterion validity, and construct validity. Here the writer focuses on content validity because content validity will show the representative of higher order thinking skills in school nationally standardized examination.

Content validity is estimated by testing the feasibility or relevance of the test content through rational analysis by a competent panel or expert judgment. Sekaran (2006) states that content validity ensures that measurements include a sufficient and representative set of items that reveal the concept. The more scale items reflect the area or the overall concept being measured, the greater the content validity. Or in other words, content validity is a function of how well the dimensions and elements of a concept have been described

Content validity is done to determine whether the contents of the test items are appropriate and relevant to the purpose of the test. Kowsalya et al (2012) argue that the content validity of the tests was obtained thoroughly and systematically in examining test items to determine to what extent the test items reflect and do not reflect the content domain. Sekaran (2006) explains that content validity shows that items intended to measure a concept is able to uncover the concept to be measured.

Surapranata (2005) states that content validity means that a test tool is valid if the content of the test appropriate with the curriculum.

To measure the content validity by comparing the curriculum with the test items that will be measured. If the whole test items or instrument used to match the curriculum, it can be concluded that test items have high validity. If there are some test items or instrument that are not appropriate with the curriculum, the content validity is relatively low. Harsiati (2012) argues that to measure the content validity is by comparing the proportion of the curriculum with the test items, comparing the suitability of the items with the curriculum and compare the scope of the test material with the scope of the curriculum.

The content validity is measured through the formula:

$$\frac{\text{Total of test items that suitable with the curriculum}}{\textit{total test items}} X\ 100\%$$

Arikunto`s criteria were used for interpreting the result of analyzing content validity.

Table 2.6 The criteria of content validity

Content Validity Index (%)	Criteria of content Validity
81-100%	Very high
61-80%	High
41-60%	Average
21-40%	Low
0-20%	Very low

Source: Arikunto (2013:89)

2.2.7.2 Reliability

A reliable test is consistent and dependable. The issue of reliability of a test may best be addressed by considering a several factors that may contribute to the unreliability of a test. Mousavi (2002) argues the following possibilities: fluctuations in the student, in scoring, in test administration, and in the test itself. Jackson (2003) states that reliability is a measure of stability or internal consistency of an instrument in measuring certain concepts. According to Creswell (2002), there are various types of reliability depending on the number of times the instruments are administered and the number of individuals who provide information.

There is a relationship between validity and reliability. The test can be reliable but not valid; however, it cannot be valid if it is not reliable. In other words, if a test is valid, it must be reliable. And, in general, checking for validity of a test is more difficult than checking for reliability because validity is measuring data related to knowledge whereas reliability only concerns with the consistency of scores.

The reliability of multiple-choice test items is measured through Kuder-Richardson KR-21.

$$r_{21} = \frac{k}{k-1}\left[1 - \frac{M(k-M)}{kS^2}\right]$$

r_{21} = reliability coefficient of the whole test

k = number of items in the test

S^2 =variance of scores

M =mean of the scores

(Rudner & Schafer, 2002, p.18)

While the reliability of the essay test items are measured through alpha Cronbach

$$r_{11} = \left[\frac{k}{k-1}\right]\left[1-\frac{\sum \sigma_b^2}{\sigma_t^2}\right]$$

r11 = Search instrument reliability coefficient

k = Number of questions or number of questions

$\Sigma\sigma_b^2$ = Amount of variance score of item i

i = 1, 2, 3, 4, ... n

σ_t^2 = Total variance

Categories of reliability can be seen in the Table 2.7

Table 2.7 the Criteria of Reliability

Reliability Coefficient interval	Description
0,810 – 1,000	Very high
0,610 – 0,800	High
0,410 – 0,600	Fair
0,210 – 0,400	Low
0,000 – 0,200	Vey low

Source: Guilford (1956, p.145)

2.2.7.3 Practicality

Test items must meet the practical aspects of understanding and implementing the test items. Practicality shows the level of ease of use and implementation which includes costs and time in the implementation and management and interpretation of the results. Therefore, the purpose of the practicality test is to find out the extent to which the ease and implementation of school nationally standardized examination test items are made.

According to Nation & Newton (2009, p.166) practicality can be observed from several aspects: (1) economy of time, money, and labor; (2) ease of administration and scoring; and (3) ease of interpretation. The school nationally standardized examination design demonstrates all the aspects. It is administered in a multiple-choice format since it is an efficient and effective way to assess a wide range of skills. It is also easier to score due to the objective assessment.

Practicality assessment based on questionnaires that have been filled in by practitioners is analyzed to determine the level of practicality of the school nationally standardized examination test items being developed. Practical analysis using a Likert Scale with steps: Giving a score for each item, very good (5), good (4), enough (3), less (2) and very less (1). Summarize the total score of each

practitioner for all indicators and providing practical value by using the formula:

$P = \frac{f}{N}$ x 100 %

P = A percentage of attainment

f = score obtained

N = Maximum Score

Categories of practicality can be seen in Table 2.8

Table 2.8 Categories of Practicality

No	Percentage (%)	Criteria
1	81-100	Very practical
2	61-80	Practical
3	41-60	Quite practical
4	21-40	Less practical
5	0-20	Not practical

Source : Riduwan (2005, p.89)

2.3 Theoretical Framework

The theoretical framework is the conceptual framework in determining the direction of research which will be implemented. It is the frame of mind from the researcher. Below is the figure which explains the theoretical framework of this study. It gives the illustration of investigating the present study.

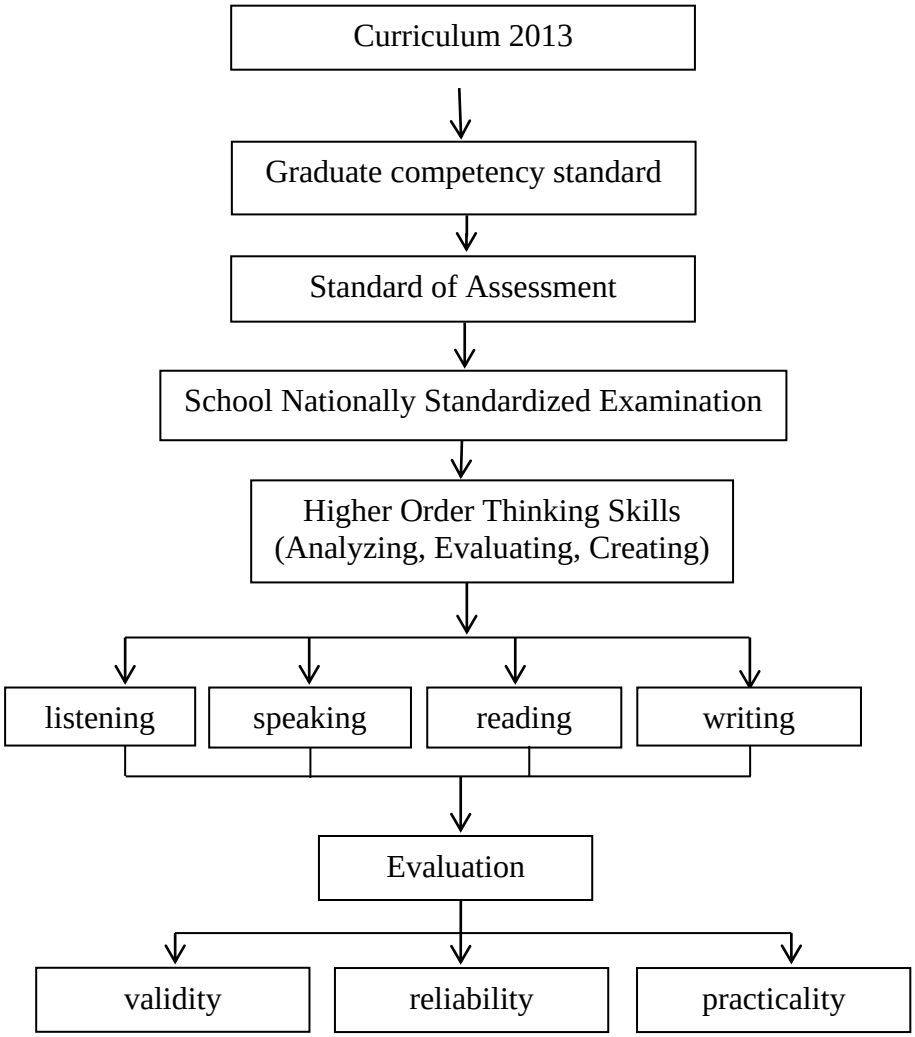

Figure 5 Theoretical Framework

This framework used to describe the evaluation of higher order thinking skills in School nationally standardized examination. From the figure above, we can see that higher order thinking skills on school nationally standardized examination clarified from some dimensions of nationally education standard. It is in the scope of curriculum 2013. Curriculum 2013 covers graduate competency standard and the assessment

standard. The assessment standard mentioned that school nationally standardized examination is the assessment conducted by the educational unit. In school nationally standardized examination there are higher order thinking skills. The researcher uses higher order thinking skills in school nationally standardized examination as the data. The data will be analyzed the implementation of higher order thinking skills in English skills. The researcher also evaluates the content validity, reliability, and practicality in English school nationally standardized examination.

CHAPTER V

CONCLUSION AND SUGGESTIONS

This chapter discussed the conclusion for this present study and suggestion for the English teacher and the next researcher.

5.1 Conclusion

This study focused on the evaluation of higher order thinking skills in English school nationally standardized examination at senior high school 6 Semarang. Based on the aforementioned results, this study concludes some important points.

The realization of higher order thinking skills in listening is realized in one question it means that the variation level of the higher order thinking needed to be enriched. As a consequence, the student`s lost their opportunity to use higher order thinking to find out the answers. However, the development of the HOTS questions in listening is needed to be enriched as they have not reached the highest level of creating.

In speaking questions the realization of higher order thinking skills does not appear, it makes the students lost their opportunity to develop their communication competence and communicative activity like transmitting information, ideas, or feeling. It noted that speaking as communication competency is one of the frameworks in 21st-century skills.

The realization of higher order thinking skills in reading questions indicate the presence of almost all levels of cognitive domains in reading questions, except “creating” which is the highest level of thinking in Bloom`s

revised taxonomy. Lower order thinking skills (LOTS) are the main concern in reading questions. The few distributions of the HOTS happen because the questions look much more complex than the LOTS. Therefore the test maker concern about the limitation of time provided for the student to work. This finding also reveals that there is still much room for reading questions be the driving force in the effort to make learners critical thinkers.

Higher order thinking skills in writing questions consist of two types namely multiple choice and essay form. Essay form in writing questions is the best way to realize higher order thinking skills because students could express the ability of critical and creative thinking, to find out and analyzing information in the problem-solving problem faced. In essay form, the questions of higher order thinking skills realized in the highest level of Bloom`s revised taxonomy creating level. HOTS questions in writing act as a guide for the students to expand their ideas on what to write. Thus when students are occupied with extra information, they will be able to extend the length of their writing as well as their perspectives and knowledge.

Overall the English school nationally standardized examination questions are suitable among the contents of the syllabus measured by determining the material in the basic competencies and the contents of English school nationally standardized examination items. These results indicate that the content validity of the questions (41 items or 91.11%) belong to very high criteria.

The reliability coefficient of the test scores in multiple-choice is 0.27 which is categorized low. While the reliability of the test score in the essay is 0.121 which is categorized very low (unreliable). From the results of the English school nationally standardized examination question reliability test, it is known that the learning outcome test has very low reliability. It means that the English school nationally standardized examination questions are not steady, not fixed, or can be said to be unreliable.

The practicality is observed from several aspects: (1) economicality; (2) ease of administration and scoring; and (3) ease of interpretation. The English school nationally standardized examination demonstrates all the aspects. It is administered in a multiple-choice format and essay format since it is an effective and efficient way to assess a wide range of skills.

Finally, from the whole results and discussions, this present study had shown that the higher order thinking skills have been realized in English school nationally standardized examination. There is still much space for English school nationally standardized examination to implement higher order thinking skills test items in the effort to make students critical thinkers.

5.2 Suggestion

In this part, the researcher presents the suggestions related to the evaluation of English school nationally standardized examination or other language tests in general.

It is recommended that the test makers should modify the proportion variation of HOTS questions in English school nationally standardized

examination. The test makers should consider the anxiety and worries of the student's failure in working the test. Then create alignment between the English school nationally standardized examination HOTS questions with the curriculum to ensure the reduction of LOTS level questions and increase the questions requiring analyzing, evaluating, and creating which belong to HOTS.

Further study needs to conduct studies related to the current one in another English school nationally standardized examination questions include to what extent the higher order thinking skills were more reflected. To ensure students succeed and prepare them to face the challenges in 21st-century, it is important to train them to have creative and critical thinking. One of the ways to reach the purpose is by providing them intensive exercises to answer questions requiring their higher order thinking skills such as those belonging to analyzing, evaluating and creating level. In this case, assessment and evaluation practices of teachers are of great importance.

REFERENCES

Abdullah, A., Alberta, S. W., & Ardiansyah. (2018). Profil of chemical teacher candidates higher order thinking skills on ionic equilibrium in solution topic. *Unnes Science Education Journal*, *7*(3), 300-305.

Abosalem, Y. (2016). Assessment techniques and students' higher-order thinking skills. *International Journal of Secondary Education*, *4*(1), 1-11.

Ahmad, U. L. (2016). Senior high school english national examination and thinking skills. *Beyond Words, 4*(2), 168-190.

Airasian, P. W., & Russel, M. K. (2008). *Classroom assessment: concepts and applications 6th Ed.* Boston: McGraw-Hill.

Anasy, Z. (2016). Higher order thinking skills (HOTS) in reading excercise. *TARBIYA: Journal of Education in Muslim Society*, *3*(1), 51-63.

Anderson, et al. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. A bridged edition*. New York: Addision Wesley Longman.Inc.

Anderson, L. W., & Krathwohl, D. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Anisah, & Lastuti, S. (2018). Pengembangan bahan ajar berbasis HOTS untuk meningkatkan kemampuan pemecahan masalah matematis mahasiswa. *Kreano Jurnal Matematika Kreatif-Inova*tif, *9*(2), 191-197.

Arikunto, S. (2013). *Prosedur penelitian: suatu pendekatan praktik*. Jakarta: Rineka Cipta.

Ariyana, Y., Pudjiastuti, A., Bestary, R., & Zamroni. (2018). *Buku pegangan pembelajaran berorientasi pada keterampilan berpikir tingkat tinggi.* Jakarta: Direktorat jenderal guru dan tenaga kependidikan kementerian pendidikan dan kebudayaan.

Armbruster, B., & Ostertag, J. (1993). *Learning from textbooks* : *theory and practice. (B. K. Britton, A. Woodward, & M. Brinkley, Eds.)* New Jersey: Lawrence Erlbaum Associates Inc.

Asrivi, Q. E. S., Rokhman, F., & Deliana, S. M. (2017). Penerapan standar kompetensi lulusan mata pelajaran bahasa Indonesia sekolah dasar. *Journal of Primary Education, 6*(3), 257-266.

Assaly, I. R., & Smadi, O. M. (2015). Using Bloom's taxonomy to evaluate the cognitive levels of master class textbook's questions. *English Language Teaching*, *8*(5), 100-110.

Bachman, L. F. (1990). *Fundamental considerations in language testing.* Oxford: Oxford University Press.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.

Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. Virginia: ACSD Member Book.

Brown, H. D. (2004). *Language assessment principles and classroom practices*. USA: Longman.

Brown, J., & Rodgers, T. (2002). *Doing second language research*. Oxford: Oxford University Press.

BSNP. (2006). *Permendiknas RI no 22 tahun 2006 tentang standar isi untuk satuan pendidikan dasar dan menegah*. Jakarta: Badan Standar Nasional Pendidikan.

Creswell, J. W., & Clark, V. L. P. (2006). *Designing and conducting mixed methods research* (1st ed.). Thousand Oaks, CA: Sage Publications.

Creswell, J. W. (2003). *Research design: qualitative, quantitative, and mixed methods approaches (2nd Edition)*, Thousand Oaks, CA: Sage Publications.

Darling-Hammond, L. (2014). *Next generation assessment: Moving beyond the bubble test to support 21st century learning*. San Fransisco: Jossey-Bass.

Destianingsih, A. (2016). Designing english learning worksheets for reading and writing skills qualifying on HOTS for the tenth grade students. *INOVISH JOURNAL*, *1*(2), 11-30.

Elyana, Yennita, & Fakhruddin. (2017). Analysis of higher order thinking skill student MAN 2 model Pekanbaru in problem solving of physics national exam. *Jurnal online mahasiswa Fakultas Keguruan dan Ilmu Pendidikan Universitas Riau*, *4*(1), 1-9.

Emmert, P. *(*1994*). A definition of listening*. Listening Post. Ad.

Fahim, M., & Komijani, A. (2010). Critical thinking ability, L2 vocabulary knowledge, and learning strategies. *Journal of English Studies*, *1*(1), 23-38.

Fanani, A., & Kusmaharti, D. (2018). Pengembangan pembelajaran berbasis HOTS (higher order thinking skills) di sekolah dasar kelas V. *Jurnal Pendidikan Dasar*, *9*(1), 1-11.

Fanani, M. Z. (2018). Strategi pengembangan soal higher order thinking skill (HOTS) dalam kurikulum 2013. *Jurnal Edudeena IAIN Kediri*, *2*(1), 57-76.

Febrina, Usman, B., & Muslem, A. (2019). Analysis of reading comprehension questions by using revised Bloom`s taxonomy on higher order thinking skills (HOTS). *English Education Journal (EEJ), 10*(1), 1-15.

Fenton, R. (1996). Performance assessment system development. *Alaska Educational Research Journal, 2*(1), 13-22.

Fitri, Z. A. (2018). Achieving competency standards: impact of excellence program. *Jurnal Pendidikan Islam*, *4*(1), 51-62.

Ghina, A., Simatupang, T. M., & Gustomo, A. (2017). The relevancy of graduates' competencies to the effectiveness of entrepreneurship education: a case study at SBM ITB. *Indonesia Journal of Entrepreneurship Education*, *2*(1).

Grellet, P. (2004). *Developing Reading Skills*. United State of America: Cambridge University Press.

Guilford, J. P. (*1956). Fundamental statistic in psychology and education*. 3rd Ed. New York: McGraw-Hill Book Company.

Hammersley, M. (1992). *Deconstructing the qualitative-quantitative divede*. In: Brannen, Julia ed. Mixing Methods: qualitative and Quantitative Research. Aldershot: Avebury

Handayani, R., & Priatmoko, S. (2012). Pengaruh pembelajaran problem solving berorientasi HOTS terhadap hasil belajar kimia siswa kelas X. *Jurnal Inovasi Pendidikan Kimia, 6*(2), 954-965.

Harmer, J. (2007). *The practice of english language teaching*. Harlow: Longman.

Harsiati, T. (2012). *Penilaian dalam pembelajaran*. Malang: UM press.

Heong, Y. M., Othman, W. D., Yunos, J., Kiong, T. T., Hassan, R., & Mohamad, M. M. (2011). The level of Marzano higher order thinking skills among technical education students. *International Journal of Social and humanity*, *1*(2), 121- 125.

Herkusumo, A. P. (2011). Penyetaraan ujian akhir sekolah berstandar nasional (UASBN) dengan teori t tes klasik. *Jurnal Pendidikan dan Kebudayaan, 17*(4), 455-471.

Hidayati, A. (2015). Prosiding Semiar Nasional Pendidikan Ekonomi & Bisnis 2015: *Relevansi kompetensi lulusan sekolah menengah kejuruan dengan kebutuhan dunia usaha dan industri.*

Hornby, A. S. (2006). *Oxford advanced learner's dictionary*. Oxford: Oxford University Press.

Hughes, A. (1989). *Testing for language teachers*. Cambridge and New York: Cambridge University

Iffa, U., Fakhruddin, & Yennita. (2017). Analysis of higher order thinking skills student in problem solving of physics science national examination. *Jurnal Online Mahasiswa (JOM) Fakultas Keguruan dan Ilmu Pendidikan (FKIP) Universitas Riau, 4*(1), 1-9.

Istiyono, E. D., Mardapi, & Suparno. (2014). Proceeding International conference on educational researchand evaluation (ICERE) 2014: *Effectiveness of the reasoned objective choice test to measure higher order thinking skills in physics implementing curriculum 2013*.

Jackson, D. N. (2003). The relative validity of scales prepared by native item writers and those based on empirical methods of personality scale construction. *Educational and Psychological Measurement*, *35*, 361–370.

Jaedun, A., Hariyanto, V. L., & Nuryadin, E. R. (2014). An evaluation of the implementation of curriculum 2013 at the building construction department of vocational high schoos in Yogyakarta. *Journal of Education*, *7*(1), 14-22.

Jailani, A, B. (2015). Proceeding of Internationally Conference On Research, Implementation And Education Of Mathematics And Sciences 2015: *Developing an assessment instrument of higher order thinking skills (HOTS) in mathematics for junior high school grade VIII semester I.*

Julianingsih, S., Rosidin, U., & Wahyudi, I. (2017). Pengembangan instrumen assesmen HOTS untuk mengukur dimensi pengetahuan IPA siswa di SMP. *Jurnal Pembelajaran Fisika Universitas Lampung, 5*(3), 59-68.

King, F. J., Goodson, L., & Rohani, F. (2013). *Higher order thinking skills.* Retrieved from http://www.cala.fsu.edu/files/higher_order_thinking_skills.pdf January 21st 2019.

Kowsalya, D. N., Lakshmi, H. V., & Suresh, K. P. (2012). Development and validation of a scale to assess emotional maturity in mild intellectually disabled children language in India.

Krathwohl, D. R. (2002) . *Revision of Bloom`s Taxonomy: An Overview, Theory Into Practice*, *41*(4), 212-218, DOI:10.1207/s15430421tip4104_2.

Kristianingsih, D. D., Wijayati, N., & Sudarmin. (2016). Pengembangan LKS fisika bermuatan generik sains untuk meningkatkan higher order thinking skills siswa. *Journal of Innovative Science Education, 5*(1), 73-82.

Kusindang, & Bharathi, D. A. L. (2016). Improving students` motivation in speaking ability by using story telling. *English Eduational Jurnal, 6*(2), 80-86.

Laily, N. R., & Wisudawati, A. W. (2015). Analisis soal tipe higher order thinking skill dalam soal UN kimia SMA rayon B tahun 2012/2013. *Kaunia, 11*(1), 27-39.

Larson, L. C., & Miller, T. (2011). 21st century skills: prepare students for the future. *Kappa Delta Pi Record, 47*(3), 121-123, DOI: 10.1080/00228958.2011.10516575.

Mahfuzah, A., Jufri, & Fitrawati. (2019). Students' ability in answering reading questions with HOTS at SMA N 3 Bukittinggi. *Journal of English Language Teaching, 8*(1), 71-81.

Martalyna, W., Isnarto, & Asikin, M. (2018). Students' mathematical literacy based on their self-efficacy by discovery learning with higher order thinking skills-oriented. *Unnes Journal of Mathematics Education Research, 7*(1), 54-60.

Marzano, R. J., & Kendall, J. S. (2007). *The new taxonomy of educational objectives*. California:Corwin Press, Thousand Oaks.

Mayangsari, D., & Santoso, R. S. (2016). Analisis kebijakan implementasi kurikulum 2013 di SMA Kesatrian I kota Semarang. *Indonesian Journal of Public Policy and Management Review, 5*(3).

Mayer, R. E. (2002). Theory into practice: rote versus meaningful learning. *The H.W. Wilson Company, 4*(1).

McCowan, R. J., & McCowan, S. C. (1999). *Item analysis for criterion-referenced tests*. New York: Center for Development of Human Services, 39. Retrieved from https://archive.org/details/ERIC_ED501716. Accessed on February 10th 2019.

Mohammadi, E. N. & Zare, Z. (2015). The relationship between critical thinking ability and listening comprehension ability of Iranian EFL

learners. *International Journal of Research Studies in Language Learning*, *4*(1), 47-59.

Mousavi, S. A. (2002). *An encyclopedic dictionary of language testing* (3rd ed.). Taipei: Tung Hua Book Company

Mulyadi, & Suprayekti. (2013). Meningkatkan efektivitas strategi pembelajaran PAIKEM terhadap nilai ujian akhir sekolah. *Perspektif Ilmu Pendidikan*, *27*(2), 89-96.

Mulyaningsih, I., & Itaristanti. (2018). Pembelajaran bermuatan HOTS di jurusan tadris bahasa Indonesia. *Indonesian Language Education and Literature*, *4*(1), 114 – 128.

Mulyasa. (2005). *Menjadi guru profesional, menciptakan pembelajaran kreatif dan menyenangkan*. Bandung: PT. Remaja Roosdakarya.

Musial, D., Nieminen, G., Thomas, J., & Burke, K. (2009*). Foundations of meaningful educational assessment.* New York: McGraw-Hill.

Musrikah. (2018). Higher order thinking skill untuk anak sekolah dasar dalam pembelajaran matematika. *Martabat: Jurnal Perempuan dan Anak, 2*(2), 340-360.

Mustapa, K. (2014). Pengaruh strategi pembelajaran terhadap kemampuan berpikir tingkat tinggi. *Jurnal Pendidikan Humaniora*, *2*(4), 248-257.

Muzayanah, U. (2015). Item quality of Islamic education test on the final exam of national standard schools. *Jurnal SmaRT, 1*(1), 125-135.

Nababan, B. O., Wau, Y., & Purba, S. (2017). Evaluation of implementation of curriculum 2013, in public senior high school 7 Binjai, North Sumatera. *International Journal of Sciences: Basic and Applied Research, 35*(1), 28-36.

Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. New York :Routledge.

Nation, I. S. P., & Macalister, J. (2010). *Language curriculum and design*. New York: Taylor and Francis Group.

National Research Council. (1987). *Education and learning to think*. Washington D.C: National academy press.

Ngongo, P. B., Relmasira, S. C., & Hardini, A. T. A. (2018). Penerapan HOTS dalam model pembelajaran small group discussion untuk meningkatkan hasil belajar PPKn. *Journal for Lesson and Learning Studies, 1*(3), 196-203.

Nunan, D. (1993). *Syllabus design*. Oxford: Oxford University Press.

Nunan, D. (2003). *Practical English languange* teaching. New York: Mc Graw-Hill.

Nuraini, S., Distrik, I. W., & Suana, W. (2018). Pengembangan lembar kerja siswa (LKS) blended learning berorientasi pada higher order thinking skills (HOTS). *Journal of Physics and Science Learning, 2*(1), 69-77.

Otaya, L. G. (2014). Indonesia standardized examination in islamic religious education. *Jurnal Pendidikan Islam UIN Sunan Gunung Djati*, *1*(2), 223-246

Peraturan Pemerintah Republik Indonesia Nomor 32 Tahun 2013 tentang standar nasional pendidikan.

Prastiwi, A., Sriyono., Nurhidayati. (2016). Pengembangan modul fisika berbasis masalah untuk meningkatan high order thinking skills (HOT) siswa SMA. *Jurnal Radiasi, 9*(1), 1-6.

Pratiwi, H. P., Hidayah, N., & Martiana, A. (2017). Pengembangan modul mata kuliah penilaian pembelajaran sosiologi berorientasi HOTS. *Cakrawala Pendidikan*, *36*(2), 201-209.

Prihantoro, C. R. (2015). The perspective of curriculum in Indonesia on environmental education. *International Journal of Research Studies in Education, 4*(1), 77-83.

Pusat Penilaian Pendidikan. (2018). *Panduan penyusunan soal ujian sekolah berstandar nasional*. Jakarta: Pusat penilaian pendidikan badan penelitian dan pengembangan kementrian pendidikan dan kebudayaan.

Qomariyah. (2014). Kesiapan guru dalam menghadapi implementasi kurikulum 2013. *Jurnal Pendidikan Ekonomi IKIP Veteran Semarang, 2*(1), 21-35.

Rapih, S., & Sutaryadi. (2018). Perpektif guru sekolah dasar terhadap higher order thinking skills (HOTS): pemahaman, penerapan dan hambatan.

Premiere Educandum: Jurnal Pendidikan Dasar dan Pembelajaran, 8(1), 78 – 87.

Riadi, A. (2016). Problem-based learning meningkatkan higher-order thinking skills siswa kelas VIII SMPN 1 Daha Utara dan SMPN 2 Daha Utara. *Math Didactic: Jurnal Pendidikan Matematika*, *2*(3), 154-163.

Riduwan. (2009). *Skala pengukuran variabel-variabel penelitian*. Bandung: Alfabeta.

Rudner, L., & Schafer, W. (2002). *What teachers need to know about assessment*. Washington, DC: Nationally Education Association.

Rumahlatua, D., Huliselana, E. K., & Takariaa, J. (2016). An analysis of the readiness and implementation of 2013 curriculum in the west part of Seram district, Maluku Province, Indonesia. *International Journal of Environmental & Science Education, 11*(12), 5662-5675.

Sadler, D. R. (2012). Assessment, evaluation and quality assurance: implications for integrity in reporting academic achievement in higher education. *Education Inquiry, 3*(2), 201–216.

Sari, J. I., Fitri, S. G. S., & Hendriyani, M. E. (2018). Pengembangan soal berfikir tingkat tinggi pembelajaran bioteknologi yang berorientasi pada pemanfaatan potensi sumber daya hayati pulau Tunda. *Biodidaktika: Jurnal Biologi dan Pembelajarannya*, *13*(1), 37-46.

Sekaran, U. (2006). *Metode penelitian bisnis*. Jakarta: Salemba Empat.

Setiawan, A., & Bharati, D. A. L. (2018). Developing HOT project-based-speaking assessment to stimulate the students' critical thinking and creativity. *English Education Journal*, *8*(3), 301-307.

Setyarini, S., Muslim, A. B., Rukmini, D., Yuliasri, I., & Mujianto, Y. (2018). Thinking critically while storytelling: improving children's HOTS and English oral competence. *Indonesian Journal of Applied Linguistics*, *8*(1), 189-197.

Setyasih, A. D., Hartono, & Prasetyo, A. P. B. (2019). Effort to increase scientific literacy skills and the thinking process of higher order thinking skills. *Journal of Primary Education, 8*(2), 144-151.

Slameto. (2018). Penerapan cooperative learning untuk meningkatkan HOTS. *PerKhasa: Jurnal Pendidikan Dasar*, *4*(1), 124-140.

Sofyan, F. A. (2018). Implementasi HOTS (higher order thingking skill) pada kurikulum 2013. *Jurnal Inventa*, *3*(1), 1-17.

Stufflebeam, D. J., & Shinkfield, A. J. (2007). *Evaluation theory, models, & applications.* San Francisco, CA: Wiley.

Subadar. (2017). Penguatan pendidikan karakter berbasis higher order thinking skills. *Jurnal Pedagogik*, *4*(1), 81-93.

Sumaryanto. (2018). Penilaian HOTS pada pembelajaran matematika. *Indonesian Digital Journal of Mathematics and Education, 8*(8), 500-509.

Surapranata, S. (2005). *Analisis, validitas, reliabilitas, dan interpretasi hasil tes*. Bandung: PT Remaja Rosdakarya.

Suwandi. (2013). Evaluasi pelaksanaan ujian akhir sekolah berstandar nasional (UASBN). *Jurnal Pendidikan Teknologi dan Kejuruan*, *21*(3), 227-237.

Syafa'ah H. K., & Handayani, L. (2015). Pengembangkan instrumen metacognitive self–assessment untuk mengukur keterampilan berpikir evaluasi dalam membaca teks sains berbahasa Inggris. *Unnes Physic Education Journal*, *4*(1), 43-48.

The regulation of the Minister of Education and Culture number 20, 2016 about gradute competency standard

The regulation of the Minister of Education and Culture number 23, 2016 about assessment standard

The regulation of the Minister of Education and Culture number 24, 2018 about National examination and school nationally standardized examination

The regulation of the Minister of Education and Culture number 59, 2014 about curriculum 2013 for senior high school

Thomas, A., & Thorne, G. (2009). How to increase higher order thinking: Retrieved from http://www.cdl.org/articles/how-to-increase-high-orderthinking/ January 21st 2019.

Thoyyibah, N., Hartono, R., & Bharati, D. A. L. (2019). The implementation of character education in the English teaching learning using 2013 curriculum. *English Education Journal, 9*(2), 254-266.

Toyoda, E. (2018). Assessment of higher order thinking skills required for intercultural learning. *Intercultural communication studies, 27*(1), 1-20.

Trilling, B., & Fadel, C. (2009). *21st Century Skills: Learning for Life in Our Times*. San Fransisco: Jossey-Bass.

Undang-undang No 20 tahun 2003 tentang sistem pendidikan nasional Indonesia. Jakarta.

Valette, R. M. (1967). *Modern language testing: a handbook*. NewYork: Harcourt, Brace, & World

Vethamani, M. E. (2008). Reading literary text. *The English Teacher*, *36* (1), 20-33.

Wahid, H. A., & Karimah, R. A. (2018). Integrasi higher order thinking skills dengan model creative problem solving. *MODELING: Jurnal Program Studi PGMI, 5*(1), 82-98.

Wardhani, Y. I. (2018). Pembelajaran thinking actively in social contexts untuk meningkatkan kemampuan higher order thinking siswa SMA. *Journal of Biology Education*, *1*(1), 1-16.

Widana et al. (2018). Higher order thinking skills assessment towards critical thinking on mathematics lesson. *International Journal of Social Sciences and Humanities*, *2*(1), 24-32.

Widana, I, W. (2017). Higher order thinking skills assessment. *Journal of Indonesian Student Assessment and Evaluation*, *3*(1), 32-44.

Widana, W. I. (2017). *Modul penyusunan soal high order thinking skill (HOTS*). Jakarta: Direktorat pembinaan sekolah menegah atas direktorat jendral pendidikan dasar dan menengah, kementrian pendidikan dan kebudayaan.

Widdowson, H. G. (1996). *Teaching language as communication*. NewYork: Oxford University.

Widihastuti. (2013). Strategi pendidikan karakter di perguruan tinggi melalui penerapan assesssment for learning. *Jurnal Pendidikan Karakter*, *3*(1), 38-52.

Widiyanto. (2013). Strategy of increasing the competence of graduates of vocational high school (SMK) majoring in business and management based business and industrialized. *IOSR Journal of Research & Method in Education, 3*(6), 25-30.

Widodo, T., & Kadarwati, S. (2013). Higher order thinking berbasis pemecahan masalah untuk meningkatkan hasil belajar berorientasi pembentukan karakter. *Cakrawala Pendidikan, 32*(1), 161-171.

Wijayanti, Y. R. (2015). An evaluation model of problem based learner assessment in curriculum 2013. *Indonesian journal of English Language Studies, 1*(2), 207-209.

Winarti et al. (2015). Prosiding Internationally Conference on Mathematics, Science, and Education 2015: *Analysis of higher order thinking skills content of physics examinations in madrasah aliyah.*

Yanto, H., Fam, S. F., Baroroh, N., & Jati, K. W. (2018). Graduates accounting competenceies in global business: perceptions of Indonesian practicioners and academics. *Academy of Accounting and Financial Studies Journal*, *22*(3), 1-17.

Yulianto, A., Hasyim, A., & Nurmalisa, Y. (2014). Analisis kesiapan guru dalam mengimplementasikan kurikulum 2013. *Jurnal Kultur Demokrasi*, *2*(5), 1-11.

Yuniar, M., Rakhmat, C., & Saepulrohman, A. (2015). Analisis higher order tinking skills pada soal objektif tes dalam mata pelajaran ilmu pengetahuan sosial (IPS) kelas V SD Negeri 7 Ciamis. *Jurnal PEDADIDAKTIKA: Jurnal Ilmiah Pendidikan Guru Sekolah Dasar, 2*(2), 187-195.