How to Construct TOS (*Table of Specifications*) for testing language?

This book is designed to be used as a reference for teachers with skills and knowledge about constructing a table of specification (TOS) as a blueprint for a test. TOS is a tool to help the teacher decide to create test construction and improve the validity of the teacher-made test for evaluation. It is also helpful for teachers to determine the kinds of examinations they need to administer based on the course materials. It involves the content material to be covered and the relative emphasis to be placed based on the materials and instructional objectives stated in the curriculum. The contents of this book were taken from various relevant and up-to-date data source which covers some chapter including testing, evaluating, and assessment in language learning, bloom taxonomy, testing by using HOTS (higher order thinking skills) and constructing test specification.

UNIPMA Press
WE GOT IT

Penerbit UNIPMA Press
Universitas PGRI Madiun
Jl. Setia Budi No. 85 Madiun, Jawa Timur 63118
E-Mail: upress@unipma.ac.id
Website: kwu.unipma.ac.id

ISBN 978-623-8095-35-3

How to Construct TOS (Table of Specifications) for testing language?

UNIPMA Press
WE GOT IT

How to Construct TOS (*Table of Specifications*) for testing language?

Nuri Ati Ningsih
Rengganis Siwi Amumpuni

How to Construct TOS (Table of Specifications) for Testing Language?

Nuri Ati Ningsih
Rengganis Siwi Amumpuni

UNIPMA Press
WE GOT IT

How to Construct TOS (Table of Specifications) for Testing Language?

Penulis:
Nuri Ati Ningsih
Rengganis Siwi Amumpuni

Editor:
Nuri Ati Ningsih

Perancang Sampul & Letak
Tim Kreatif Unipma Press

Cetakan Pertama, November 2023

Diterbitkan Oleh:
UNIPMA Press Universitas PGRI Madiun
JI. Setiabudi No. 85 Madiun Jawa Timur 63118
E-Mail: upress@unipma.ac.id
Website: kwu.unipma.ac.id
Anggota IKAPI: No. 207/Anggota Luar Biasa/JTI/2018

ISBN: 978-623-8095-35-3

Preface

First of all, let us express thanks to Allah. The Almighty, for providing us with grace and benefits to accomplish this book entitled "How to Construct TOS (Table of Specifications) for Testing Language?". This book is designed to be used as a reference to teachers about constructing a table of specification (TOS) as a blueprint of a test that involves the content material to be covered and the relative emphasis to be placed based on the materials and instructional objectives stated in the curriculum. TOS is a tool to help the teacher decide to create test construction and improve the validity of the teacher-made test for evaluation. It also helps the teacher to identify the types of tests they need based on the course of materials.

The contents of this book were taken from various relevant and up-to-date data source which covers some chapter including testing, evaluating, and assessment in language learning, bloom taxonomy, testing by using HOTS (higher order thinking skills) and constructing test specification.

The writers would like to express their gratitude to Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) Universitas PGRI Madiun which sponsored the program, the Head of the English Department, and the Rector of Universitas PGRI Madiun who allowed the writers to write this book. In addition, a big thanks to all the contributors who helped accomplish this book

The Writers

Contents

Chapter 1
Testing, Evaluation, and Assessment

A. Testing

A.1 What is Testing?

The definition of the word testing provided by the Oxford Dictionary is as follows: (noun) the activity or process of testing or checking someone or something; (adjective) showing a person's capabilities by placing them under pressure; difficult. The root meaning of the word "test" is "to take steps to check the quality, performance, or reliability of (something)," and this is done primarily prior to bringing something into general application or practice. A test is a method that functions to determine the student's ability, knowledge, or performance in a certain domain. It could be in the cognitive, affective, or psychomotor domain. Brown (1994) proposes that a test is a procedure that is carried out in order to evaluate a student's capability to carry out tasks in a given area within a certain period of time in order to achieve a specified objective. According to Le Grange & Reddy (1998), A test is a tool, technique, or approach that is meant to examine students' knowledge or their ability to do a certain activity. Tests might come in the for of multiple-choice questions, essay questions, or performance-based questions. When viewed in this light, testing might be understood to be a type of

assessment.Therefore, it can be concluded that a test refers to a tool, technique, or method to measure the student's ability, knowledge, and performance. In other words, it is used to measure students` abilities or competencies.

A.2 Types of Testing

Brown (1994) classified testing into several types. They are;

a. Placement tests

   A test of this kind is intended to provide an educator with guidance regarding the placement of a student into a specific level or section of a language curriculum or school.

b. Diagnostic tests

   This test helps teachers and learners to assess the student's linguistic knowledge and language skills before a course is begun.

c. Proficiency tests

   The purpose of this examination is to determine the student's level of language. It is not tied to any specific method of education but rather evaluates the learner's overall level of language knowledge.

d. Language Aptitude Test

   This test is designed to evaluate a student's potential for acquiring linguistic knowledge. This test consists

of several different test items which measure such abilities as:

1. Sound- coding ability

i.e the ability to identify and remember new sounds in a new language

2. Grammar-coding ability

i.e the ability to identify the grammatical functions of different parts of sentences

3. Inductive learning ability

i.e the ability to work out meanings without explanation in the new language

4. Memorization

i.e the ability to remember and to recall words, patterns, and rules in the new language

e. Achievement tests

This test is intended to measure the students' skills and knowledge after they do some kind of instruction given by their teachers.

A.3 Classification of Testing Task

A.3.1 Matching Test

The matching test is almost similar to multiple choice. It has similar characteristics to multiple-choice tests, but it also has

different. One of them is each option can be matched to one of several stems.

Examples include

a. matching words or sentences to the pictures they describe
b. matching the words they use to their meanings
c. matching sentence fragments or matching responses to their questions
d. matching dialogues or texts to likely contexts
e. selecting the most appropriate headings or summary statements to correspond with the individual paragraphs or parts of a text

These words are important for understanding the writing process. Match each word with the correct definition.

a. step	**1.** to check a piece of writing for errors
b. topic	**2.** a group of related sentences
c. gather	**3.** one thing in a series of things you do
d. organise	**4.** subject; what the piece of writing is about
e. paragraph	**5.** to change or correct a piece of writing
f. essay	**6.** a short piece of writing, at least three paragraphs long
g. proofread	**7.** to arrange in a clear, logical way
h. edit	**8.** to find and collect together

Source: http://penilaiangrammar.blogspot.com/2009/07/simple-completion-sentences.html

A.3.2 Multiple Choice Test

The multiple-choice item test format is a type of flexible test format that can be used to assess the integration of ideas stated in the texts. In a language test, it involves vocabulary, grammar, and propositional meaning.

Example:

26. Phillip's _____ tone endeared him to his comical friends, but irritated his serious father.
- **a.** aloof
- **b.** jesting
- **c.** grave
- **d.** earnest
- **e.** conservative

27. Brian's pale Irish skin was _____ to burn if he spent too much time in the sun.
- **a.** prone
- **b.** urbane
- **c.** eminent
- **d.** erect
- **e.** daunted

Source: Pustaka Bhs Inggris 2017

A.3.3 True False Test

The explicit meaning of individual phrases or units of ideas in spoken language can be evaluated through the use of tests referred to as true/false items. In another way, it is also used to measure the student's ability to judge the accuracy of summaries of opinions expressed in a certain text.

Example:

I have a kitten named Simba, he is a 4 month old domestic cat. We named him after the character in the movie The Lion King because of his color. Simba is a little bit chubby but he likes to play and jump around everywhere. If he sees something interesting, like keyrings, or anything else that is hanging, he will jump and try to get it. Sometimes, if he is not in the mood to play, he will sit cutely somewhere. He also loves to snuggle with me and I love him for that.

28. What is the kitten's name?

a. **Simba**

b. Aurora

c. Kiko

d. Poo

Source: http://penilaiangrammar.blogspot.com/2009/07/simple-completion-sentences.html

A.3.4 Information Gap

This type of task gives the students incomplete information. Students must try hard to combine their knowledge to complete a task.

The type of information gap includes:

- Differences between the pictures
- Completing the table and the picture
- Putting together a story from its constituent parts

Example:

Look at the picture !

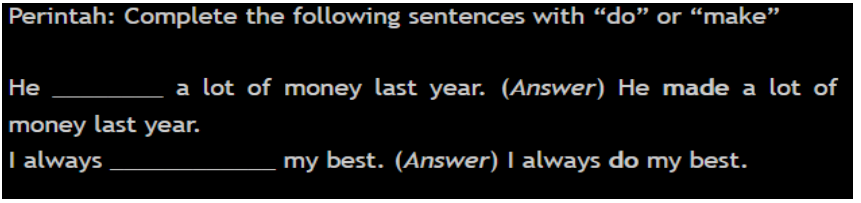
Are there two vases in your classroom?. . .
○A. Yes, there are
○B. No, there are not
○C. Yes, there is
○D. No, there is not

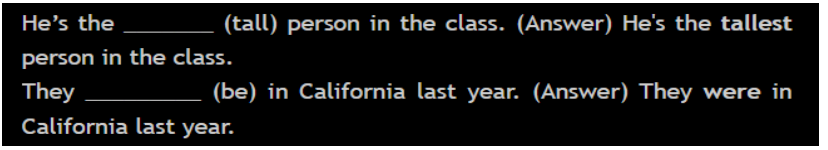
Source: https://www.diaryguru.com/2021/08/soal-bahasa-inggris-things-in-the-classroom.html

A.3.5 Completion Test

The completion questions and tests are a sort of exam that ask the student to finish an incomplete statement that is missing one or more significant components like words, sentences, numbers, and so on. It is generally agreed that completion questions are the most effective way to test for memorization of important information and are also superior for evaluating lesser levels of cognitive ability.

Example:





sentences.html

A.4 Developing Test

According to CEFR (2011),

Some stages must be done by the teachers in administering the test. They are described in the following figure;

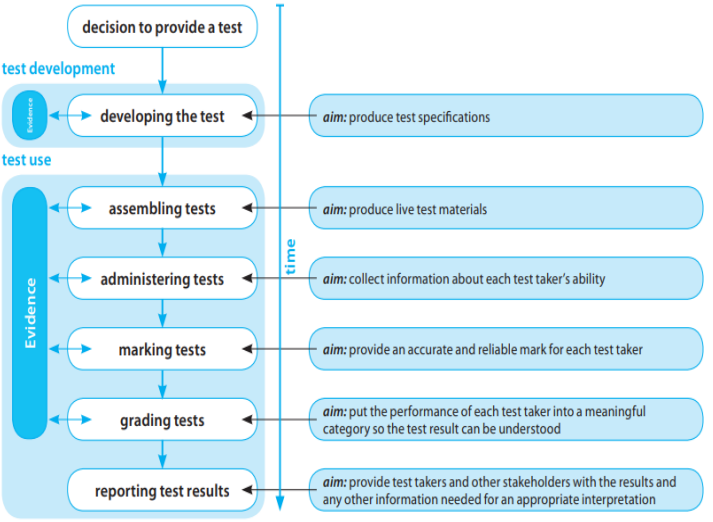


Figure 1.1 The basic testing cycle

B. Evaluation

B.1 What is Evaluation?

Thakur, P proposed some definitions of evaluation based on some experts. They are;

Thorndike and Hegan:

Evaluation is a term closely related to measurement. In some cases, it includes informal and intuitive judgment of the students learning progress. Evaluation describes the thing related to a certain component that is chosen and judges the student's degree of acceptability or suitability of the material given.

James M. Bradfield:

Evaluation is the assignment of symbols to the phenomenon, to characterize the worth or value of a phenomenon, usually concerning some cultural or scientific standards.

Norman E. Gronlund and Robert L. Linn:

The term "evaluation" refers to a methodical procedure that entails gathering, evaluating, and interpreting information in order to assess how well students achieve educational goals (https: //www.yourarticlelibrary.com/education/evaluation-in-education-meaning- principles-and-functions/89656).

Evaluation has a wider meaning covering academic and non-academic achievement the students have gained. In short, evaluation can be defined as the process of collecting, analyzing, and interpreting the information about the students done systematically and it has a purpose to determine the extent to which students have achieved the instructional objectives. There are some types of evaluation. Detail types and descriptions of each type of evaluation are described as the following;

a. Formative Evaluation

   Formative evaluation is a type of evaluation that is conducted during the process of teaching. The purpose of this type is to provide continuous feedback to the teacher and the students. It is to evaluate the students after completing or finishing a lesson or topic. The result of this

evaluation is to get an idea of whether they have to change the teaching method or not.

b. Summative Evaluation

Summative evaluation is a type of evaluation conducted at the end of an academic year/certain period. The purpose is to evaluate the achievements of instructional objectives and the students' change in the overall personality at the end of the period. The summative evaluation also considers a comprehensive aspect of learning. This evaluation not only takes into account the grade of formative evaluation but also tests the students after completing a certain course to get the final grade and give feedback to the students. Grading is conducted based on summative evaluation.

c. Prognostic Evaluation

Prognostic evaluation is used to estimate and predict the students' future careers. A prognostic evaluation broadens the result of an assessment by analyzing the abilities and potential with the further dimension, the future development, the necessary condition, the time frame, and the limit of the person.

d. Diagnostic Evaluation

A diagnostic evaluation is used by the teacher to evaluate the student's strengths, weaknesses, knowledge, and skills

before the beginning of the teaching-learning process. The result of this type of evaluation is used to design the course and the curriculum. Conducting the diagnostic test needs special preparation and techniques.

e. Norm-Referenced Evaluation

This type of evaluation is done by comparing the students' performance individually within (1) the group of students being tested, (2) the group of students of similar age, (3) the group of students of similar experience, and (4) the group of students of similar background. The result of the evaluation will determine the position of students in the group.

f. Criterion-Referenced Evaluation

This type of evaluation represents the individual performance of the students compared with a fixed standard performance. It portrays the accuracy of the student's performance, such as; how well the individual performance concerning a specific standard.

g. Quantitative Evaluation

Quantitative evaluations are conducted by using scientific tools and measurements It is possible to quantify or count the findings of this evaluation. Oral tests, written tests, and performance tests are all examples of the kinds of evaluation methods that might be utilized for this purpose.

h. Qualitative Evaluation

The characteristics of this type of evaluation are more subjective than quantitative evaluation. Qualitative evaluation is established in science based on the use of the five senses. It involves value judgment, Then, the techniques or tools used in Qualitative Evaluation are Cumulative Records and Anecdotal records (https://www.your article library. com/education/ evaluation-in-education-meaning-principles-an function /89656).

B.2 The Purpose of Evaluation

a. Improvement of Learning

Giving evaluations to the students will contribute directly to the development of the students in the learning process. The result of an evaluation gives real feedback to the teacher about the improvement of the student's ability and their progress in learning.

Continuous evaluation provides the instructor with information regarding the level of learning at each stage. If the students find difficulties or gaps in their learning, an appropriate remediation can be provided to overcome this situation. For the students who show good progress in their learning, enrichment measures can be initiated. Therefore,

evaluation contributes to the improvement of learning through the processes of diagnosis and remediation.

b. Improvement in Teaching

The results of the evaluations can determine whether the poor performance of the students is caused by poor teaching, ineffective teaching methods, the absence of the teachers, or the teachers' ignorance in the teaching process. Thus, evaluation can be used as an important instrument for improvement in teaching.

c. Renewal of Curriculum or Course Content

The outcome of the assessment provides comments and details about the relevance and usefulness of the material covered in the course. Modification of the curriculum can be done if we find unsuitable material or content for the students. Specific information is very useful in defining the learning objectives as well. So, evaluation can provide a basis for curriculum revision.

d. Development of Non-Cognitive Capacities.

In this era, the main concern of education is not enough on intellectual development. Teachers also have to develop other aspects such as social intelligence, emotional intelligence, and physical aspects of personality, and the vital one to be developed today is mental intelligence. So, comprehensive evaluation takes into all the aspects above

in it (https://targetb-ed.co.in/what-is-evaluation-in-education-types-of-evaluation assessment -for- learning-b-ed-2nd-year/).

B.3 The Characteristics of Good Evaluation

a. Evaluation is an Objective-based Process

Evaluation is used to know the development of the student's personality and their achievement in the process of teaching and learning. In this case, both are reflected in the teaching-learning outcome or the instructional objectives. Meaningful evaluation must be in tune with these instructional objectives/ learning outcomes. Instructional objectives offer some directions used for the teaching and evaluation process. The technique and tool of evaluation also must be selected based on the characteristics of instructional objectives. So, evaluation is done to measure or to see whether the instructional objectives have been achieved or to what extent the instructional objectives have been done.

b. Evaluation is a Continuous Process

The continuous process of evaluation means that evaluation is essential for getting reliable evidence about the students ' growth and development of their teaching-learning process. As a result of this, the objective of evaluation needs to be integrated with teaching so that it

can assist teachers in correctly diagnosing the challenges that their students are having and so that it can offer opportunities for students to receive remedial instruction. The students' improvement in learning will not be done without the process of continuous evaluation. In short, evaluation should not be considered as an end-of-the-course or learning activity.

c. Evaluation is a Comprehensive Process

Every student has different dimensions of growth, such as intellectual, emotional, and physical aspects. These aspects will be represented in the form of different objectives. A comprehensive evaluation must provide all the aspects of the process of evaluation.

d. Evaluation is a Cooperative Process

This characteristic is related to a comprehensive process. Comprehensive evaluations have to collect evidence on all aspects of the student's development. In this case, the teacher cannot collect all the evidence of all aspects required for the student's growth or development. It deals with the student's social relationships, emotional behavior, initiative, scientific: attitudes, social attitudes, and likes. Accordingly, a good evaluation program must have the cooperation of different individuals and agencies in collecting the data.

e. Evaluation is a Dynamic Process

Evaluation is conducted by the teachers based on instructional objectives. It also helps them to know whether the instructional objectives are compatible with a certain group of students or not. The evaluation is conducted on the basis of the learning experience that has been established in the classroom; in addition, it gives proof of the efficiency of their learning experience. Regular feedback in evaluation maintains the validity of the whole process of teaching and learning. So, a good evaluation program will bring a dynamic process and it will lead to continuous improvement of evaluation in the whole process of education.

f. Evaluation is a Decision-Making Process

The appraisal process is needed in every step of teaching and learning. In the first step (before the instruction), it is necessary to decide the appropriate instructional objectives, learning materials, teaching-learning strategies, and learning material. The second step (during instruction) continuous evaluation is needed to measure the student's intellectual, emotional, and physical aspects to make a proper decision by considering various remediation programs. Both diagnostic and formative evaluation must go hand in hand to achieve the objectives. The last at the

end of the session is giving a summative evaluation. This is done so that the students can be categorized, graded, promoted, and certified. As a result, the teachers will be better able to judge and decide on the educational progress of the students with the assistance of summative evaluation.

B.4 The Principle of Evaluation

The process of evaluation has to be carried out with effective techniques. To get an effective one, the teachers have to do the following steps;

a. It must be clearly stated what is to be evaluated

   First, a teacher has to establish the purpose of evaluation. Second, he/ she must determine the instructional objectives/learning outcome, and third, define them clearly in terms of the student's observable behavior.

b. A variety of evaluation techniques should be used for a comprehensive evaluation:

   To get a comprehensive evaluation, teachers have to evaluate all the aspects of achievement. Some techniques can be used e.g. objective tests, essay tests, observational techniques, etc. Comprehensive evaluation represents the complete picture of the student's achievement and the student's development can be assessed extensively.

c. An evaluator should know the limitations of different evaluation techniques:

Teachers have to know well about the characteristics of every technique and instrument of evaluation. It involves the strengths and the weaknesses of each of them. There may be measurement errors. Sampling error is a common factor in educational and psychological measurements. Measurement errors can also be found due to students guessing on objective tests. Error is also found due to incorrect interpretation of test scores.

d. The technique of evaluation must be appropriate for the characteristics or performance to be measured:

Every method of evaluation is suitable for certain applications, but there are situations in which it is unsuitable for others. Therefore, in order for teachers to properly select an evaluation method, they need to have a solid understanding of both the benefits and drawbacks of the methods that are available to them.

e. Evaluation is a means to an end but not an end in itself:

The result of the evaluation will be used to make decisions about the students. It is not just collecting the data about the students. Teachers have to consider the data of all aspects of evaluation and take them in making judgments.

B.5 Evaluation in Teaching and Learning Process

Teaching and learning process without evaluation is not possible. Evaluation is a part of the teaching and learning process. Both teaching and learning are done based on the instructional objective or learning outcomes that present some directions to reach them. Instructional objectives or learning outcomes are set up based on some desirable students' behaviour to be developed along the process of teaching. The relationship between the instructional objectives, the process of teaching/learning experience, and evaluation can be described in the following diagram;

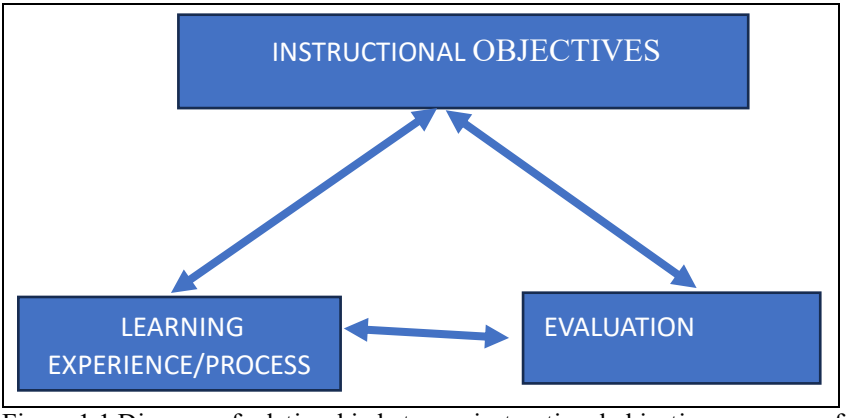


Figure 1.1 Diagram of relationship between instructional objectives, process of learning experience, and evaluation

C. Assessment

C.1 What is assessment?

In the teaching process, an assessment plays an important role. The results of this assessment are very useful for the teacher to find out the students' progress, achievement, and even career prospects in the future. Assessment can be conducted informally during the instruction (teaching and learning process) or at the end of instruction, usually, it is called a formal assessment. Assessment is the procedure or process of gathering and synthesizing information from tasks (such as exams on performance or learning) in order to create an opinion about a person or to make a comparison against a predetermined standard (Athanasou and Lamprianou: 2002). According to Susanti, et al (2021), Assessment is a collection and processing process of information to determine learning needs, development, and achievement of student learning outcomes.

C.2 The Function of Assessment

Types of assessment according to their function include assessment as a learning process (assessment as Learning), assessment for the learning process (assessment for Learning), and assessment at the end of the learning process

The tendency in the implementation of the previous assessment	The implementation of the expected assessment
	

Figure 1.2. The Change of The Implementation of Assessment that gives the emphasis on Formative Assessment

The purpose of conducting an assessment in education is to determine the student's progress in learning and to make curriculum decisions or revisions

Assessment as a learning process has a function as a reflection of the learning process and as a formative assessment. Assessment for learning has the purpose of revising the learning process and as a formative assessment. Assessment of learning means assessment done at the end of the process of learning as evaluation and has functioned as a summative assessment. The difference between assessment and learning is assessment as learning involves students more actively in the assessment activity. Students are provided with the opportunity to gain the experience necessary to understand how to be an objective evaluator of both themselves and their peers. Examples of evaluation as learning include both self-assessment and assessment conducted by peers on peers. In assessment as

learning, students should be involved in formulating procedures, criteria, and rubrics/guidelines assessment so they know exactly what must be done to obtain maximum learning outcomes (Susanti, et al: 2021).

Furthermore, Susanti et al (2021) explain that one example of implementing formative assessment is self-assessment and assessment between friends (peer assessment). This assessment works as material for self-reflection, which can later be used by Educators as data/information to confirm student learning outcomes.

Table 1: The Example of Self-Assessment

Give a mark on self-assessment of competency in writing descriptive text. How is your understanding?	
	I easily wrote it
√	I can write it
	I need help to write it

Table 2: The Example of Peer-Assessment

Task: Writing Descriptive Text Appraiser's name: …… The name of a friend being rated: …. Give a sign on the appropriate one!	
√	The message conveyed is clear.

	The grammar used doesn't match with the descriptive text.
√	The generic structure of the text doesn't meet the requirement of descriptive text (identification and description)

In order to the implementation of the assessment is in line with the objectives to be achieved, educators are expected to pay attention characteristics and functions of formative and summative assessments. Susanti et al (2021) explain in detail about formative and summative as classified in several tables below :

Table 3: Formative (as and for learning)

Types of Assessment	**Formative (as and for learning)**
Function	a. Diagnosing abilities beginning and learning needs learners. b. Feedback for educators to improve the process of learning to become more meaningful. c. Feedback for the students to improve learning strategies. d. Diagnosing the capacity of the students to absorb the material along the process of learning in class. e. Driving the change of class atmosphere. It can improve the students' motivation by providing supportive, positive, and meaningful learning programs.

Technique	Various techniques of assessment, such as; practice, product, project, portfolio, and written test/ oral.
Result or Documentation	a. The product of learning b. The students' reflection journal c. The plan of the next action is designed based on the result of the assessment. d. The record or the note based on the result of observation e. Anecdotal notes f. The score in the form of number.

Table 3 : Formative (as and for learning)

Types of Assessment	**Formative (as and for learning)**
Function	f. Diagnosing abilities beginning and learning needs learners. g. Feedback for educators to improve the process of learning to become more meaningful. h. Feedback for the students to improve learning strategies. i. Diagnosing the capacity of the students to absorb the material along the process of learning in class. j. Driving the change of class atmosphere. It can improve the students' motivation by providing supportive, positive, and meaningful learning programs.
Technique	Various techniques of assessment, such as; practice, product, project, portfolio, and written test/ oral.
Result or Documentation	g. The product of learning h. The students' reflection journal

	i. The plan of the next action is designed based on the result of the assessment. j. The record or the note based on the result of observation k. Anecdotal notes l. The score in the form of number.

Table 4: Summative (for and of learning)

Types of Assessment	**Summative at the end of the material (for and of learning)**
Function	a. Measuring tools to find othe ut achievement of learning outcomes students in one scope of material. b. Reflection on learning in one material scope. c. Feedback for design/repair of the next learning process. d. Seeing or finding the power/strengthen, and weaknesses of learning on students during one-scope learning material.
Techniques	Various techniques of assessment, such as; practice, product, project, portfolio, and written test/ oral.
Result or Documentation	1. The product of learning 2. The score in the form of number

Table 5: Summative Semester (of Learning)

Types of Assessment	**Summative Semester (of learning)** ***As a choice**
Function	a. Measuring tools to find out the student's achievement of learning outcomes in a certain period. b. Getting the value on the achievement of learning outcomes to be compared with the criteria of achievement that have been set.

	c. Feedback to design/repair the next learning process (semester/school year) the same as the function of formative assessment d. See/ find the power and the weaknesses of the students in the learning process (same as the function in the diagnostic assessment)
Techniques	Practices, products, : projects, portfolios, written.
Result or Documentation	1. The product of learning 2. The score in the form of number

According to Henriani, S, and Suzzane (2013), assessments are given to find the answers to some questions, such as:

a) How realistic are my teaching plans for these students?
b) Are my students ready for the next unit?
c) What learning difficulties are students facing?
d) Which students are underachieving?
e) How effective was my teaching? Which learners are advanced?
f) Which learners are gifted or talented?
g) Which learners require special assistance?

The assessment is one of the aspects of education. The result of it will help the teachers to make judgments about students' levels, to choose the most appropriate teaching method, and to determine the teaching of a topic. Based on the description above, it can be concluded that assessment is the method of process in, collecting,

documenting, and using empirical data related to the students' knowledge, skills, attitudes, and beliefs. These data will be used to increase and improve the process of student learning.

An education test is used by teachers to examine the student's knowledge of something, to determine what the student knows or has learned. The test measures the students' level of skill or knowledge that has been reached. A test is a kind of "product" that is used to measure a particular student's behavior or set of objectives determined by the teacher in the teaching and learning process. On the contrary, the assessment is supposed as a procedure instead of a product.

Assessment is implemented during and after the instruction has taken place. After receiving the result of the assessment, we can interpret the result and if it is needed, we can change the teachers' instructions. Tests are done after the instruction has taken place to complete the instruction and get the results of it. The results of the tests don't have to be interpreted, unlike the result of the assessment. ssessment is a continuing process that encompasses a significantly larger subject area or body of content. On the other hand, testing is just one component of the assessment process. They are predetermined procedures that are carried out at specific periods throughout a program of study. To evaluate students, teachers employ a variety of methods and activities, including tests, which can be categorized as helpful assessment

devices. It can be concluded that a test is a type of method used to measure a person's ability, knowledge, or performance in a given domain.

Assessment can be divided into:

a. Informal and Formal Assessment

Informal assessment is a type of assessment that can be given in a number of forms. It can be started with incidental time, by giving unplanned comments and responses, along with coaching and other types of feedback given to the student.

Example: "Nice job!"

"Good work!"

Formal assessments are given in the form of exercises or procedures. It's specifically designed to measure the students' skills and knowledge. The characteristics of the formal assessment are systematic, planned sampling techniques, and it is constructed to give teacher and student an appraisal of student achievement. If there is a question, such as "Is formal assessment the same as a test? We can say that all tests are formal assessments, but not all formal assessment is testing."

b. Formative and Summative Assessment

Most of our classroom assessment is classified into formative assessment. The purpose is to evaluate the

students in the process of “constructing” their competencies and skills based on the teacher's instructions to help them to continue that growth process. A summative assessment is carried out after the conclusion of a unit of instruction or a course in order to evaluate or summarize what a student has achieved throughout the learning process. Generally speaking, summative assessments take place at the end of a course. Examples of summative evaluation include a course's final examination as well as general competence examinations.

The principle in conducting an assessment involves;

1. Assessment is an integral component of the learning process, serving to both facilitate and direct learning while also giving educators, students, and parents with comprehensive information in the form of feedback that can assist them in formulating a plan for additional learning activities
2. The assessment is designed and carried out by the assessment function, with the freedom to determine the technique and time for assessing to effectively achieve learning objectives
3. Assessments are designed to be fair, proportional, valid, and can be trusted (reliable) to explain learning progress and determine decision about the next step
4. Reports on the development of learning and the accomplishments of participants. Education is uncomplicated and

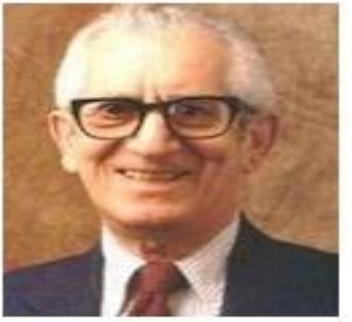
imparts a wealth of knowledge, delivering important details regarding the personalities and skill sets developed as well as the next steps to take.

5. The results of the assessment are utilized as resources for reflection to improve the overall quality of learning by the students, educators, and other members of the educational staff as well as by the parents.

Chapter 2

Taxonomy Bloom

A. Who is Bloom?



Benjamin S. Bloom

Benjamin Samuel Bloom is known as the founder of Taxonomy Bloom. Bloom was born on 21 February 1913 in Lansford, Pennsylvania. His father's profession was a picture frame and his mother was a housewife. He has three older brothers and one younger sister. Bloom has high potential capabilities in the academic field. Besides, he also very potential in athletics such as swimming and handball. He earned a bachelor's and master's degree from Pennsylvania State University in 1935. Then, he worked as a research worker with the Pennsylvania State Relief Organization for a year. After that, he moved to Washington and had the same profession in the American Youth Council. In 1939, Bloom then applied for his doctorate program at the University of Chicago. Between the years 1940 and 1959, he was a member of

a Board of Examinations at the University of Chicago. By the year 1960, Bloom had resigned from his position with the Board of Examinations and began working at Stanford University in California for the Advanced Study in Behavioral Sciences. Bloom was very loyal to his small family (his wife Sophie, two sons, and his nieces and nephew).

Bloom was outstanding as an American educational psychologist. He was pointed as headed a group of cognitive psychologists at the University of Chicago. Bloom gave significant contributions to both fields, those are the classification of educational objectives and the theory of mastery learning. Based on the result of his research, Bloom devised a "taxonomy of educational objectives," which was a classification of the various learning objectives and skills that educators set for their students. Bloom classified educational goals as falling into one of three categories, or "domains": affective, psychomotor, or cognitive. In the educational environment, this classification of educational objectives is commonly used to determine the different levels of learners' mental competencies and to dig up them from lower to higher-order thinking levels. His classification of educational objectives in the form of *Taxonomy of Educational Objectives was written in the form of a special book known as Handbook 1: Cognitive Domain which* was published in 1956. This book addresses the cognitive domain

versus the psychomotor and affective domains of knowledge. The following picture describes the old version of Bloom's taxonomy.

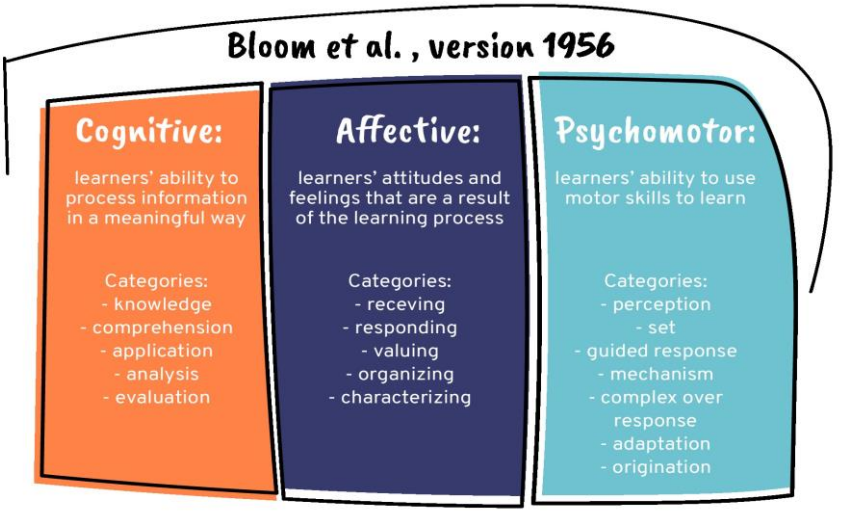


Figure 2.1: The old version of Bloom Taxonomy
https://gosiapytel83.net/objectives-taxonomies-101-part-1-of-4-blooms-taxonomy-revision/

B. Cognitive Domain

The cognitive domain was described by Benjamin Bloom as one of the best-known educational domains. It is related to the process of the development of our mental (thinking skills and the process of acquiring our knowledge. This domain is focused on intellectual skills such as critical thinking, problem-solving, and creating a knowledge base. The cognitive domain covers six

categories which involve knowledge, comprehension, application, analysis, synthesis, and evaluation.

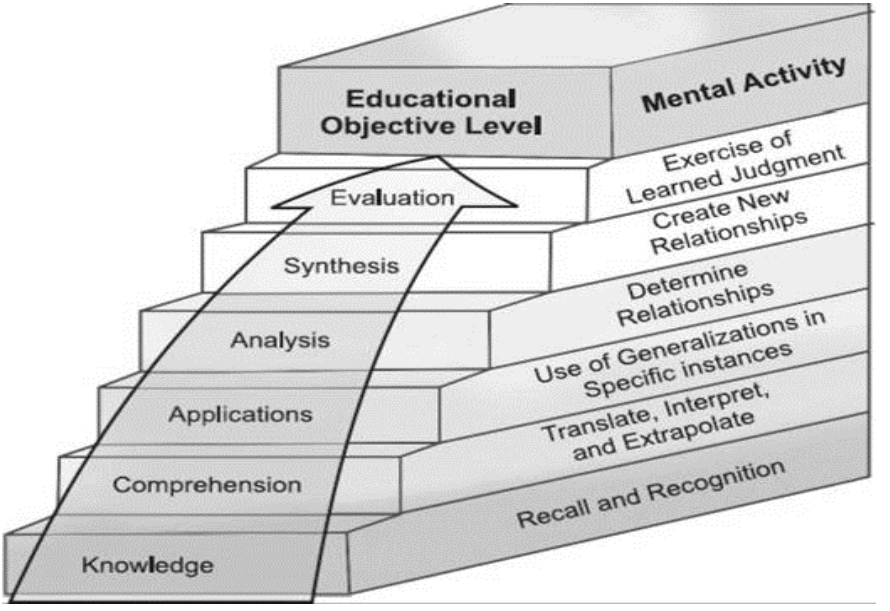


Figure 2.2 Cognitive Domain
https://ebrary.net/2967/management/basic_levels_learning_domains_learning

The description of the hierarchy above starts from the lowest level "knowledge" which emphasizes developing the ability of the students' memorization and ends on the level of the ability to evaluate the things that they have learned.

Some experts cognitive psychologists, curriculum theorists instructional researchers, and testing and assessment specialists (Lorin Anderson and Krathwohl) revised Bloom's Taxonomy in

2001. David Krathwohl was Bloom's colleague and Lorin Anderson was a former student of Bloom. The revision was done to be more adaptive to the development of the current era by proposing another taxonomy that will meet curriculum designers, teachers, and the student's needs. Hopefully, it was better than Bloom's one. The result of the revision of Bloom's Taxonomy was given the title *A Taxonomy for Teaching, Learning, and Assessment*.

What is the difference between Bloom's version and Anderson's version of taxonomy?

Bloom's version uses nouns in categorizing the level meanwhile the Anderson version uses verbs for each level. It will influence the way we demonstrate these abilities as things we perform. The highest form of learning based on the Anderson taxonomy puts "creativity" at the top level of the taxonomy. The main difference between both is more useful and comprehensive dealing with the way taxonomy acts upon the different types of knowledge, conceptual, procedural, and metacognitive. The detailed revision is described in the following picture;

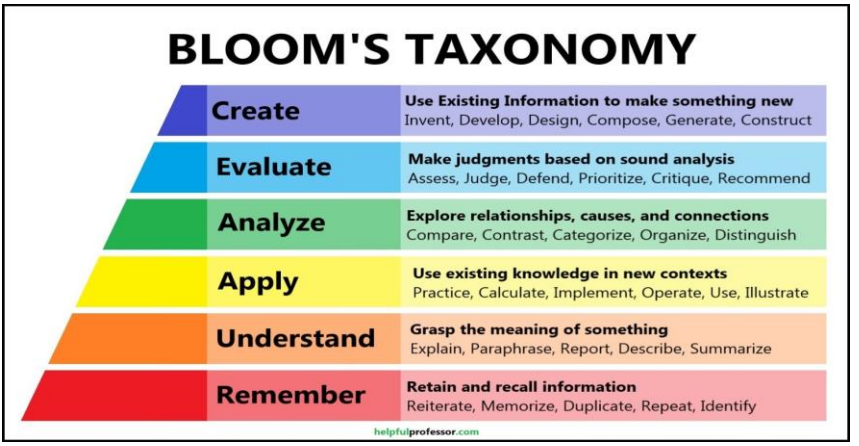


Figure 2.3 Cognitive domain revision
https://helpfulprofessor.com/levels-of-understanding/

In this new version, the cognitive hierarchy above increases from the simple ability of memorization. It was not only intended to increase the knowledge of learners, but also to generate a new idea in the learners' mental skills based on the information that they had already learned. In this instance, it is anticipated that the students will go in steps, beginning with "remember" and concluding with "create" something new.

C. Affective Domain

The affective domain is one of the domains having by taxonomy bloom. According to Krathwohl, Bloom, and Masia, (1973), the affective domain involves how dealing with emotional intelligence, such as feelings, values, appreciation, enthusiasms, motivations, and attitudes. Anderson, et al (2011)

state that the affective domain contains the feelings, emotions, and attitudes of an individual. They classified the domain into several classifications. These include receiving phenomena, responding to phenomena, placing a value on phenomena, organizing phenomena, and characterizing phenomena. The hierarchy of the affective domain is described in the following picture;

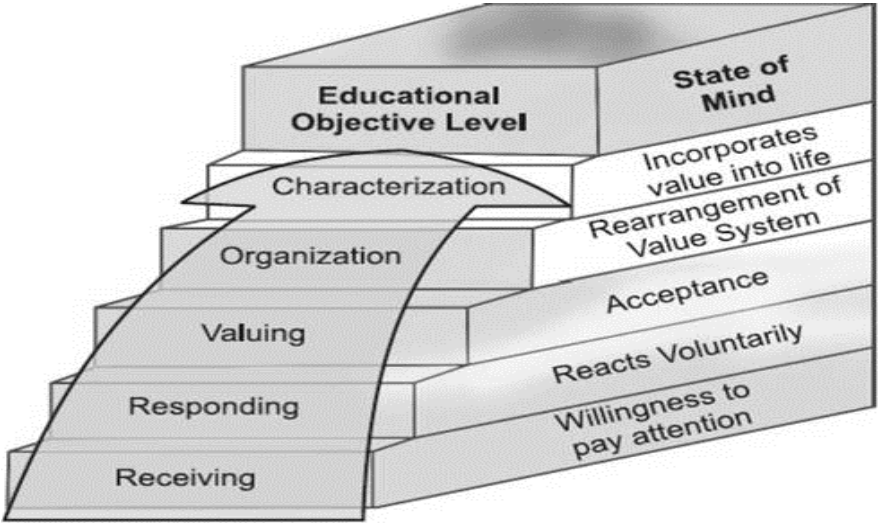


Figure 2.4. Affective Domain
https://ebrary.net/2967/management/basic_levels_learning_domains_learning

The affective domain is hierarchically constructed from the lowest level up to the highest one. Like a cognitive domain, the higher level of its construction is more complex and depends on the student's mastery on the lower level. Achieving and measuring the educational objective in this domain is not an easy thing because

it deals with the student's emotional intelligence, such as the student's attitude, personal beliefs, and values. Cannon and Feinstein (2005) explain in detail the characteristics of each level or subdomain.

1. Receiving phenomena will create an awareness of the feelings and emotions of individuals. It also supports the ability to employ the selected attention. Real class activity, include the students' activity in listening to the lesson attentively, listening to other respectively, and listening for and remembering the name of the newly introduced people.

Figure 2.5. Receiving level

2. Responding to phenomena concerns with the students' commitment to their environment. Some activities that represent this level involve the student's participation in class during the process of teaching, giving presentations in class, and giving

questions to the new concepts, models, and ideas to fully understand about it.

Figure 2.6 Responding level

3. The level of valuing covers the students' ability to see the worth of something and the ability to express it. It includes the ability to share their idea or opinion on some issues raised in classroom activity, to show the ability to solve problems, and the ability to propose a plan for a certain social environment.

Figure 2.7 Valuing level

4. Organization is the next level of the affective domain. At this level, the students can prioritize a value over another value. Hopefully, the students are also able to create a unique value system, create a life plan in harmony with their abilities, interests, and beliefs, prioritize time effectively to meet the needs of the organization, family, and self, and accept respectively for one's behaviour.

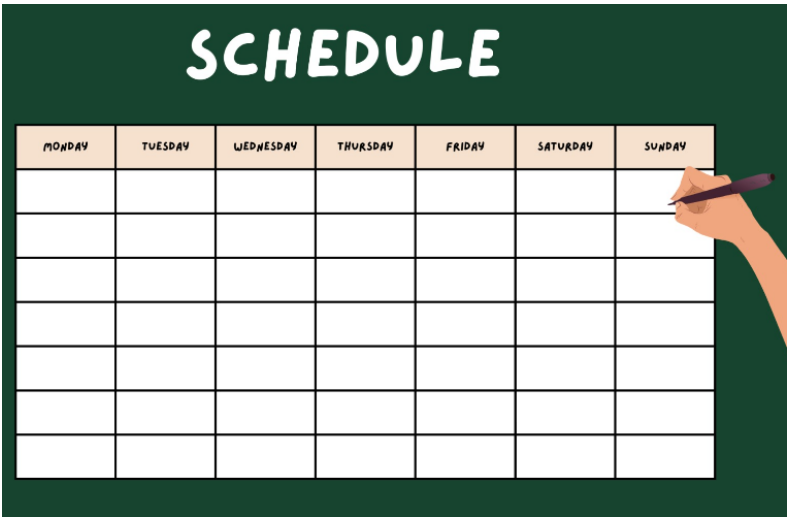


Figure 2.8. Organization level

5. The highest subdomain having by affective domain is characterization. It describes the students' ability to internalize the values and use them to control the behaviour of the individual.

Figure 2.9 Characterization level

D. Psychomotor Domain

According to Simpson (1972) psychomotor involves the physical movement, physical coordination and the use of motor-skill areas. Developing these skills needs physical practice. Measuring these skills can be done in terms of speed, precision, distance, procedure, or techniques in doing the exercises or practices. So, the psychomotor skill ranges from manual tasks to more complex ones. Dave (1970) and Harrow (1972) propose the subdomain/level of psychomotor into several levels.

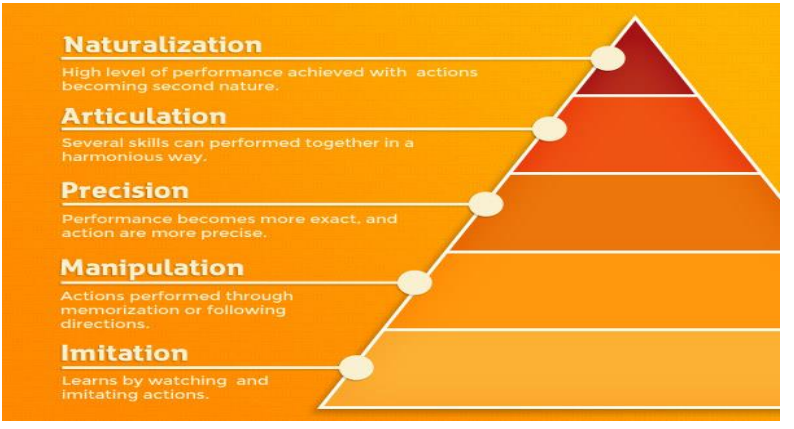


Figure 2.10. Characterization level
https://www.vectorsolutions.com/resources/blogs/teaching-skills-the-psychomotor-domain-of-learning-and-learning-objectives/

Dave (1975) classified the level of psychomotor domain as the following;

1. Imitation

After observing someone's activities, the students can copy or imitate their performance although in low quality. The classroom activities which determine this level involve the English teacher ask the students to read some words given by giving the direction "repeat after me", and the dance teacher giving directions to their students to imitate his/her performance by "follow me".

Figure 2.11. Imitation level

2. Manipulation

The students are able to do a certain or an additional action based on their memory or the teachers' instruction.

Figure 2.12. Manipulation level

3. Precision

This performance needs some actions done independently based on the instruction given orally, visually, or even in a written form. The students are expected to be able to duplicate an action with self-control. It aims to minimize an error that happened in doing an action. The real ability of the students at this level is the beginner students can perform the tasks given by the teacher for example drawing or dancing.

Figure 2.13. Precision level

4. Articulation

At the articulation level, the students need the ability to coordinate a series of actions by establishing the appropriate sequence and performing the acts accurately, with control as well as the speed and timing in doing a certain activity. The students' activities dealing with this level involve creating video animation and constructing an essay.

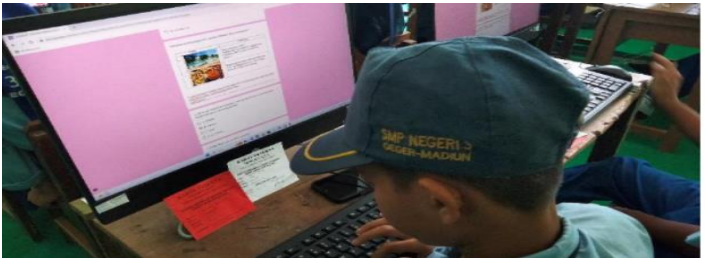


Figure 2.14. Articulation level

5. Naturalization

This level determines the students' mastery of certain skills or abilities related to physical performance, such as; the students' ability to play piano, the students' ability to operate a computer accurately and quickly, etc.

Figure 2.15. Naturalization level

Chapter 3
Testing By Using HOTS

A. What is HOTS?

HOTS stands for Higher Order Thinking Skill and it is an educational concept applied to education innovation and based on learning taxonomies. It was first found in the book "Taxonomy of Educational Objectives: The Classification of Educational Goals written by Benjamin S. Bloom, et al in 1956. Bloom defined HOTS as an abstract ability by people stated in the cognitive domain of the taxonomy of educational goals, including the ability to analyze, synthesize, and evaluate things. Then, Anderson, et al (2001) proposed a new opinion that Higher Order Thinking Skills, often known as HOTS, is sometimes described as the stage that comes after the top three stages. They are currently engaged in the process of analysis (C4), followed by evaluation (C5), which is defined as an activity involving the management of decisions and the formulation of conclusions, and finally creation (C6). This process tends to use the cognitive process more than others. Brookhart (2010) states that higher-order thinking (HOTS) is conceived of as the top end of Bloom's cognitive taxonomy. Higher-order thinking, on the other hand, is defined by Heong et al. (2011) as the practice of thinking broadly in order to discover new difficulties. In order to engage in higher-order thinking, one must first use new information or knowledge that they have

obtained and then manipulate that information in order to determine whether or not there is a possible answer to the new circumstance. Higher-order thinking skill is considered more difficult to learn or teach but also more valuable because it involves the learning of complex judgmental skills such as the process of critical thinking and problem-solving. Kuswana (2012) stated that HOTS is classified into high-level thinking which involves critical and creative thinking guided by the truth ideas and each of them has meaning. Critical and creative thinking are interdependent, as are criteria and values, reason and emotions, and so on. In detail, Ridwan Abdullah Sani describes that someone who has high-order thinking skills will be able to use their new information or knowledge to manipulate the information they get to find the solutions or to answer a new problem they have. So, it can be concluded that HOTS is considered an important skill needed by individuals to make the process of thinking innovative and creative (Ganapathy & Kaur, 2014). A detailed description of HOTS proposed by Anderson & Krathwohl (2001) is as follows;

Table 3.1: Description of HOTS

HOTS	Create (C6)	- Create your own idea - Verb: construct, design, create, develop, write, formulate
	Evaluate (C5)	- Make up your own ideas - Verb: Evaluate, judge, argue, decide, choose, support

	Analyze (C4)	- Specifying elements/aspects - Verb: compare, examine, criticize, test
MOTS	Apply (C3)	- Using the information on a different domain - Verb: Use, demonstrate, illustrate, operate
	Understanding (C2)	- Explain the idea/concepts - Verb: Explain, classify, accept, report
LOTS	Knowledge (C1)	- Recalling - Verb: remember, register, repeat, mimic.

Schraw et al. (2011) classify Bloom's thinking skills into two categories. These are Lower Order Thinking Skills which consist of knowledge, understanding, and application, and Higher Order Thinking Skills which consist of analysis, evaluation, and evaluation. The description and the keywords of each category can be seen in the following table;

Table 3.2. Description and Key Word of Bloom's Taxonomy Revision

CATEGORY	KEYWORDS	
Remembering: can the student recall or remember the information?	Mention the definition, imitate the pronunciation, state the structure, pronounce, repeat, state	LOTS-Lower Order Thinking

Understanding: Can the students explain Skill the concept, principle, law, or procedure?	Classify, describe, explain the identification, place, report, explain, translate, and paraphrase.	
Applying: Can students apply their understanding in new situations?	Choosing, demonstrating, acting, using, illustrating, interpreting, arranging schedules, making sketches, solving problems, writing	
Analyzing: can students classify the sections based on their difference and similarity?	Examining, comparing, contrasting, distinguishing, discrimination, separating, testing, doing experiments, asking	HOTS-Higher Order Thingking Skill

Evaluating: can students state good or bad about a phenomenon or certain object?	Giving argumentation, defending, stating, choosing, giving support, giving assessment, doing the evaluation	
Creating: can students create a thing or opinion?	Assemble, change, build, create, design, establish, formulate, and write.	

Nur, Ikhwan (2019) distinguished the different activities of the students when they learned in both situations LOTS and HOTS as described in the following table.

Table 3.3 Learning Differences LOTS and HOTS

Student activities in LOTS learning	Student activities in HOTS learning
Passive in thinking	Active in thinking
Problem-Solving	Formulate problems
Reviewing simple problems	Reviewing complex problems

Convergent thinking	Think divergent and develop ideas
Learning from the teacher as the main source of information	Finding information from various sources
Practice solving questions and memorizing	Critical thinking and solving problems creatively
Prioritizing factual knowledge	Analytical thinking, evaluative thinking, and making a decision

Based on some tables above, it can be simplified that the The cognitive domain is divided into six levels, which are denoted by the acronyms LOTS and HOTS. The acronym LOTS stands for long-term storage of information and the capacity to retrieve previously acquired knowledge. Knowledge, comprehension, and application are the first three levels that make up this level. In the meantime, the term "HOTS" refers to the thought process that is carried out at the most advanced levels of cognitive processing. This level is represented by the level of analysis, synthesis, and evaluation.

HOTS is essential to be implemented in the educational field to meet the challenges of the 21st century. Parents and teachers in the 21st century have a lot of tasks to realize the importance of preparing young people who are creative, able to

think critically, can make decisions properly, and are skillful at solving problems. Fadel (2008) states that HOTS is needed in the process of teaching and learning in the 21st century. The learning and innovation skills needed are creativity (creativity), critical thinking skills (critical thinking), the ability to collaborate (collaboration), and communication skills (communication).

B. The Principle of Assessing HOTS

HOTS is used to measure the ability of a high level of the student's competence. These levels go through memorizing the information caught by the students and they focus on developing the analytical skills of the students. By having analytical skills, hopefully, the students become indigenous learners and thinkers then they will be self-learners, curious learners, self-directed students, and able to find out and apply the information to solve the problem in their real life. The students with higher-order thinking skills have specific characteristics. They will have the ability to

1) relate the different concepts found in their life,
2) interpret some facts and accidents they found in their daily life
3) solve the problem they got wisely
4) communicate with the society around effectively
5) make smart negotiations with the people around, and

6) make the right decisions based on some positive considerations.

Cox, J (2019) proposes some strategies that can be used by teachers to develop the students' higher-order thinking skills. It involves;

1. Help the students determine what higher thinking skills is

   Teachers have the task of explaining to their students about higher-order thinking skills. They have to be able to encourage their students to reach this level for their thinking skills. Showing the real activities of the higher-order thinking skills implementation is also needed to make them understand well about it.

2. Connect the concepts

   Teachers have to be able to lead the students through the process of how to relate or connect one concept with another concept. This is one of the ways the teacher teach their students to be able to relate what they already know with what they are learning today.

3. Teach the students to infer

   Teachers have to be able to encourage the students to make inferences. Inference in this case refers to a kind of idea or conclusion that is drawn by the students based on some evidence or reasoning. An example of this activities is by

giving the students a certain picture of a people standing in line at a soup kitchen. Then, ask the students to look at and focus on the detailed picture. Finally, ask them to make inferences based on the picture they see.

4. Encourage questioning

   Teachers have to motivate the students to ask some questions in classroom activities. If there are no questions delivered by the students during class time, teachers can show the way how they can create the questions or keep the questions to be given for the next meeting. The class conditions must support the students to be creative. Creativity in this case refers to the students' feeling free in delivering the questions in classroom activities.

5. Use graphic organizers

   According to Fisher and Schumaker (1995), graphic organizers are a type of visual display of the essential components of information that are intended to be of assistance to students who struggle with organizing information. Its purpose is to assist students with seeing the manner in which thoughts are arranged within a particular text or in the context of a certain topic. In addition to this, it offers the students a framework for understanding more complex concepts that are presented in particular books. Students are better able to connect ideas and understand how

those ideas are related to one another when they design diagrams or mind maps.

6. Teach problem-solving strategies

   Teachers have to teach the students to use higher-order thinking skills in solving the problems they find. Using step-by-step methods for solving the problems and encouraging them to use an alternative method as well as different methods to solve the problems.

7. Encourage Creative thinking

   Teachers have to be able to lead their students to invent, imagine, and imagine what they are thinking. These activities determine that the students have creative thinking. Creative thinking helps the students to process and understand the information they get better.

8. Use mind movies

   This is a unique way in developing higher-order thinking skills. When the students find difficulties in understanding a certain concept, teachers can use this way to lead their students. Creating a movie in their mind can be done by closing their eyes and imagining it like a movie playing.

9. Teach students to elaborate on their answers.

   Students are required to comprehend an idea in order to develop higher-order thinking skills. Teachers have a responsibility to encourage students to develop their replies

by asking the appropriate questions that put them in a position where they may express their ideas in greater depth.

10. Teach QARs

    QARs is Question-Answer-Relationships. The teachers are responsible for teaching the students to identify the type of question that is being asked, and then having the students use that information to assist them in formulating an answer. It has been discovered that this particular type of method is more helpful for higher-order thinking. This is due to the fact that the students will have a greater awareness of the relationship between the material that is mentioned in a text and how they can relate it with their prior knowledge.

In assessing higher-order thinking skills, teachers have to consider some basic principles have to be done;

1. Specify clearly and exactly the kinds of thinking and content
2. Design the performance or test item that requires the students to use the target thinking and content knowledge

The example of blueprint for Senior High School Assessment.

Content Outline	Remember	Understand	Apply	Analyze	Evaluate	Create
Relation with native Americans	Identify names, dates, and events			Explain how colonial relations		

				with Native Americans were influenced by land, food and recources, political events, and the French.		
25 point, 25%	5 point, 20%			20 point, 80 5		

Adapted from Brookhart.M.Susan (2010)

Furthermore, Brookhart.M.Susan (2010) suggested three principles for assessing higher-order thinking skills. It can be done by;

(1)Using introductory material

Using introductory material means that teachers are allowing the students to use resource material or giving the students something to think about before doing something. For example;

(a)Multiple choice test

Giving introductory material first then followed by several multiple questions, and ask the students to answer them based on the material.

(b)Constructed response(essay) test

Teachers give introductory material and then ask the students to answer some questions based on the material.

(2)using novel material

Novel material in this case means that teachers set a situation that makes the students have actually to understand and think about the material deeply and they do not merely remember or just recall the material discussed in their class activities. So novel material makes the students not work with what already they have learned as a part of classroom instruction. Because of this situation, teachers should develop the students' main set with higher-order thinking skills. By setting this kind of activity, higher-order thinking skills will become a habit in their life. Dealing with novel ideas, solving problems, and thinking critically should not be something new for them that they / the students feel they "never did before". Finally, when they arrive at a summative assessment that requires the implementation of higher-order thinking skills, they will be familiar with practicing it although using other novel material.

(3)Managing cognitive complexity and difficulty separately

Naturally, the students have two concepts of thinking level. They are easy and difficult to do. Almost all of the students will suppose that recall is "easy" and higher-order thinking skill

is “hard”. To avoid this type of perception, please consider the following example;

	Easy	Difficult
Recall	Who is the main character in The Cat in the Hat?	Name all the characters in Hamlet
Higher Order Thinking	Why do you think the Cat cleaned up the house on his way out, before Mother got home?	Hamlet wrestles with a major question in his soliloquy,”(), that this too, too solid flesh would melt” In Act 1, scene 2, Lines 131-161. What is the question in his mind, and how do you think he resolves it by the end of his soliloquy? State your interpretation of his mayor question and his resolution, and use evidence from the speech to support it.

Some strategies to give feedback and scoring to assess HOTS such as;

1. **Formative assessment**

   Teachers can use observation and discussion to assess the student's higher-order thinking skills. The process of discussion can be used to observe the feedback in the form of spoken feedback (the students' reasoning) and written feedback (substantive) given based on the learning objectives.

2. **Summative assessment**

   Multiple choice question

   This type of question generally will be scored with one point for the correct answer and no point for the incorrect answer. The questions/ the steam of multiple-choice items have to be designed based on the characteristics of HOTS. Choosing the right answer among several options available in every number determines the process of thinking.

   Constructed-response and essay questions

   In giving scores for this type of task, teachers need to create scoring rubrics. The rubrics have to determine the criterion and the type of thinking to be assessed.

   Performance Assessment

   The process of scoring performance can be done based on an analytical rubric. The rubrics should figure out the quality of

thinking. Teachers can create this kind of rubric by their own criteria or by selecting a rubric from some resources (Brookhart.M.Susan: 2010).

C. Assessing Analysis, Evaluation, and Creation (HOTS)

Assessment will have a significant impact on the process of learning and teaching. For instance, it will influence the teacher in determining the material that will be given and also the method of teaching that will be used. Assessments developed based on the criteria of HOTS have some benefits, such as;

1. Evaluating the students' capacity to find a solution to a new challenge by making use of the information and skills they have acquired, explaining the phenomenon in a scientific and constructive manner, and using skills related to scientific inquiry to the process of finding a solution to the challenge.
2. Creating an environment that encourages the students to practice not only their ability in analyzing, evaluating, and creating processes (HOTS) but also to master the material they have learned.
3. Increasing the students' achievements and their motivation to learn the materials.

Some benefits above also supported by Susan M. Brookhart in 2010. She suggests that there are some benefits of HOTS assessment as the following;

1. Increase the students' motivation to learn. Teachers sometimes fail to raise the motivation of the students because the learning material given in class is still in the form of abstract and theoretical material. So, they have difficulties in understanding the material. Therefore, teachers should be able to connect the subject matter or material given in the classroom with real-world context. Assessing the students' ability by using assessment HOTS is usually given in the form of cases, natural phenomena, or contextual issues that happened in daily life. The characteristics of materials to assess the students' ability in this case relate to real-world context. So, the students will feel that their learning in the classroom is very useful to solve problems in their everyday life. This situation hopefully can increase the motivation of students to learn.
2. Improving the achievement of the students' learning outcomes. The students' motivation will strongly influence the achievement of the students.

C.1 Assessing Analysis

"Analyse, evaluate, and create" are cognitive processes represented in Bloom's taxonomy which is classified into higher-order thinking skills. Taxonomies are useful for categorizing learning objectives and assessments according to the level of complexity. Brookhart (2010) recommended to all teachers that in constructing teaching instruction and assessments should match their intended learning target in both content (what the student learns) and cognitive complexity (what the student is able to do with the learning).

Based on the taxonomy, analysis is defined as "taking information apart and exploring relationships." The type of question at the analysis level presents the students with some material they have learned (or asks them to locate the material), then ask them some questions or presents problems with answers that require differentiating or organizing the parts reasonably. Explaining the reasoning used to relate the parts to one another is often part of the analysis task.

1. Focus on a question or main idea

One central analytical skill is the ability to focus on a question or main idea, or "getting the point" of something in the form of a text. At the analysis level, focusing on the text will make the students find the main idea in it although it is not stated clearly. On the other hand, if the main idea is stated clearly in a text, students just need

to remember and understand it easily. This type of activity is classified on the lower level of the six level of Bloom's taxonomy –it just on remembering and understanding activity.

Give the students a fresh piece of introduction material at the outset so that you can evaluate how well they can focus on a topic. This could be a declaration of an issue or a policy, a political address or cartoon, the results of an experiment, a lecture, a documentary, a scenario, or a series of events. Then, after presenting students with this material, ask them to determine what the main point, problem, thesis, or argument is as a whole.

2. Analyse arguments or theses

Assessing the students' ability to analyze arguments or thesis can be done by giving the students an argument, text or a speech. Then followed by giving some questions related to the argument, text or a speech; such as

a. What evidence does the author give that supports the argument?
b. What evidence does the author give that contradicts the argument?
c. What assumptions need to hold for the argument to be valid?
d. What is the logical structure of the argument?
e. Are any parts of the statement irrelevant to the argument?

3. Compare and Contrast

We have to be careful in constructing this type of item test because not all compare and contrast item tests will measure the students' higher-order thinking skill. More complex comparison and contrast questions require the students' analysis level of thinking. An example of this type is;

Give your students a comparison and contrast task by asking them to choose two objects. They ask them to identify at least for attributes of each that were alike and four that were different, use a Venn diagram, and then write an essay (Brookhart.M.Susan : 2010).

C.2 Assessing Evaluation

Bloom stated that evaluation is used to dig up the students' ability to "critically examine information and make judgments." According to Brookhart.M. Susan (2010) to assess evaluation, teachers need to create items or tasks that direct the students to judge the values of certain materials and provide some methods for their intended purposes.

Students can appraise the material against the existing criteria. Teachers have to define the criteria in specific ways, such as standard criteria determined by the teacher (for example: literary, historical, and scientific) and the students' criteria which are proposed by the students based on their knowledge and

experience. So, this type of evaluation is done based on a reasoned evaluation that is supported by some evidence. Furthermore, in giving feedback or scoring the students' work, teachers need a scoring rubric that consists of some suggested criteria.

Following are the examples of scoring rubrics proposed by Brookhart.M. Susan (2010):

1. Criteria for feedback or rubrics
    a. Clear, appropriate statement(s)
    b. Appropriateness of evidence
    c. And soundness of reasoning and clarity of explanation
2. A holistic rubric based on some criteria

2	Completely and clearly	Uncontrolled aspects of the experiment are clearly identified and evaluated as inadequate. Reasoning is explained and is related to the concept of experimental control. An explanation is clear.
1	Partially	Some uncontrolled aspects of the experiment are identified and evaluated as inadequate. Some reasoning may not be clearly explained or not entirely related to the concept of experimental control.
0	No	Uncontrolled aspects of the experiment are not identified, or no evaluation is given. Reasoning is missing or not related to the

		concept of experimental control. Explanation is not clear.

C.3 Assessing Creation

Creation is the highest level of cognitive domain in Bloom's taxonomy. This level indicates the student's ability to create a new way of organizing things into something new. To assess the students whether can create or not, teachers can give a task to do or give them a problem to be solved. This type of evaluation has to cover the students' activity in generating multiple solutions, designing a method or procedure to achieve a certain goal, or producing something new. Those activities are classified in the level of "Synthetic" level of the old Blooms' taxonomy.

An example of assessing creation;

1. **Formative and Summative uses of results**
2. **Students self-assessment and use of result as a Special Case**

D. Assessing Creativity and Creative Thinking

According to Ker, Barbara, creativity is the ability to make or otherwise bring into existence something new, whether a new solution to a problem, a new method or device, or a new artistic object or form. The characteristics of creative people

(1) Autonomy. It means that creative people tend to be independent and non-conformist in their thoughts and actions.

(2) Creative people show an interest in apparent disorder, contradiction, and imbalance.

(3) Creative people have a high degree of self-assurance. Some possess an exceptionally deep, broad, and flexible awareness of themselves. Others are shown to be intellectual leaders with a great sensitivity to problems.

(4) Creative people can combine curiosity and problem-seeking.

(5) Creative people usually have high intelligence.

Mark Runco, an American psychologist, proposes that the creative process consists of six essential stages, or phases. They are;

1) Orientation.

It refers to a time of intense interest and curiosity. The creative people gather information in this stage.

2) Incubation.

It consists of defining the problem and seeking a solution and involves processing large amounts of information; this can occur at a conscious or an unconscious level.

3) Illumination.

This is marked by divergent thinking, openness, and excitement.

4) Verification.

In this stage, the individual evaluates his work and compares it with what is known in the field.

5) Communication.

The individual then presents his work to the industry, making it accessible to industry professionals who will evaluate its merit and potential applications.

6) Validation.

The work in this stage becomes available to society and is consequently supported or rejected.

According to the Partnership for 21st Century Skills, some requirements can be applied to be a creative person or to think creatively. The teachers can ask the students to;

1. Recognize the importance of a deep knowledge base and continually work to learn new things,
2. Open to new ideas and actively seek them out,
3. Find "source material" for ideas in a wide variety of media, people, and events,
4. Look for ways to organize and reorganize ideas into different categories and combinations, and then evaluate whether the results are interesting, new, or helpful,
5. Use trial and error when they are not sure of how to proceed, viewing failure as an opportunity to learn.

There are some activities and procedures that can be applied in classroom activities to foster the students' creative thinking.

1. Brainstorming

   Brainstorming is a type of classic creative activity to dig up the students' knowledge. All the ideas proposed by the students will be accepted and listed in this session. Evaluation will be conducted later. This approach has so many advantages. It will (1) make the number of the students' ideas rise as much as possible, (2) expose all the students in a group to share his/her ideas or opinion, (3) extend the student's thinking process, and (4) help them to see how being open to receive ideas from others to be useful.

2. Writing Reader-Response Logs in reading or literature classes

   It is a type of creative activity for the students. Teacher gives direction to the students to read a certain text or select it. After reading the text, they are asked to share the information about it, share and describe their opinion about it, sharing about their feeling, and give others types of reaction related to the text. In giving respond to this types of directions, students can do it in different types of respond, such as by connecting to some factors or by relating to their experience dealing with the content of the text. This process

of activity can be called as “reorganizing” (Brookhart, S.M.: 2010).

Assessing creativity can be conducted by using creative assignments. Creative assignments must fullfill the following rules;

1. Require the students to produce some new ideas or products or require the students to recognize the existing ideas in a new way.
2. Allow the students to make a choice (creation of an idea) on a certain matter dealing with the learning objectives to be assessed.
3. If it is graded, evaluate the students’ work against the students’ criteria to be reached.

In assessing creativity, teachers have to consider the taxonomy of creative design. It describes in detail the ways we assess the students’ creativity. One example of assessing creativity is the following;

1. Examine the picture that's currently displayed on your screen for a few second. Take note of the distinction between "imitation" and "original creation" with regard to the degree of originality displayed in both form and content.

2. Determine what level students are being creative by using the following tool:
   a. Is the new invention essentially the same as something that already existing, or is it exactly the same? This is an example of the student employing imitation. ion.
   b. For variation: Is the creation a slight change to an existing object, such that it is different, but still retains the identity of the original object?
   c. Combination: Is it a mixture of two or more things, to the point where it may be said that it is either both or all of those things?
   d. Transformation: Is it a re-creation of anything in a new context, such that it has some features of the original object, but it cannot be argued that it is still that kind of object?
   e. And finally, does it appear to have no discernible qualities of pre-existing objects or ideas? This student has created an original creation (Brookhart, S.M.: 2010).

E. The procedure to create a HOTS assessment for English Subject

Some steps must be considered by the teachers in preparing for the assessment required for HOTS (Widana, I Wayan: 2016).

1. Analyzing KD (standard competency)

It can be done by choosing or finding KKO (operational verbs) which can be categorized into HOTS classification. Please remember that **NOT** all KD can be made about HOTS. So, choose KD which requires the ability to analyze, evaluate or create not KD for remembering, understanding, and applying.

Table 3.4. Cognitive classification based on standard competency class X

No	Standard Competency	Level Cognitive
1	3.4. Membedakan fungsi sosial, struktur teks, dan unsur kebahasaan beberapa teks deskriptif lisan dan tulis dengan memberi dan meminta informasi terkait tempat wisata dan bangunan bersejarah terkenal, pendek dan sederhana, sesuai dengna konteks penggunaanya.	C-4 (Analyze)
2	4.4. teks deskriptif	C – 5 (Evaluate)

	4.4.1 menangkap makna secara kontekstual terkait fungsi sosial, struktur teks, dan unsur kebahasaan teks deskriptif, lisan dan tulis, pendek dan sederhana terkait tempat wisata dan bangunan bersejarah terkenal	
3	4.4.2 menyusun teks deskriptif lisan dan tulis, pendek dan sederhana, terkait tempat wisata dan bangunan bersejarah terkenal, dengan memperhatikan fungsi sosial, struktur teks dan unsur kebahasaan, secara benar dan sesuai konteks	C-6 (Create)

2. Writing Blueprint

A blueprint test is described as a main element of a test which includes the content or material to be tested, the amount of emphasis allocated to each area, and other important features related to testing.

Blueprints usually can be used to guide the teacher in developing HOTS assessments. It will lead the teacher to;

(1) determine the minimal standard competency that will be used to create HOTS questions

(2) choose the material related to standard competency that will be measured and formulate indicators

(4) determine the cognitive level.

Table 3.5. The Example of Blueprint

No	KD (BC)	Material	Class	Indicator	Cognitive Level	Item Type	Item number

a. Identifying the basic competence

KD. 3.5 Kelas X

Membedakan fungsi sosial, struktur teks, dan unsur kebahasaan beberapa teks khusus dalam bentuk pemberitahuan (announcement), dengan memberi dan maminta informasi terkait kegiatan sekolah, sesuai dengn konteks penggunaannya.

b. Determining the material

Besides KD, another component that should be considered is material. The material found is *"fungsi sosial, struktur teks, dan unsur kebahasaan beberapa teks khusus dalam bentuk pemberitahuan (announcement), dengan memberi dan maminta informasi terkait kegiatan sekolah, sesuai dengan konteks penggunaannya."*

c. Determining the cognitive level

We find KKO (operational verb) *"membedakan*" based on the standard competency above. This word can be used as a baseline to classify the cognitive level into C4 or the ability to analyze.

d. **Formulating indicators**

It can be done by using the ABCD formula;

A	Audience	*refers to the audience/students*
B	Behaviour	*refers to the student's behaviour that will be measured after the learning process.*
C	Condition	*refers to the condition* in which the behavior is shown by students in the learning process.
D	Degree	*it refers to the level of* appearance like what we expect from the students. We expect that the students will answer the question correctly.

The component in the indicator is the actor /audience (A) referring to the participant students, behavior (B) refers to the thinking skills to be tested, for example "can distinguish", and the prerequisite (C) refers to the existing stimulus, such as "given some text...".

So, the three elements above can be summarized in an example of an indicator statement as follows:

Example:

Given two announcement texts (C), students (A) can distinguish the social function of the two texts (B).

No	Standard Competency	Indicators
1	4.6.1 menangkap makna secara kontekstual terkait fungsi sosial,struktur teks, dan unsur kebahasaan teks pembahasan ilmiah(discussion) lisan dan tulis, terkait isu kontroversial dan actual	Disajikan sebuah data berbentuk grafik, peserta didik dapat menentukan ide utama dan menuliskan bukti pendukung ide utama dengan kata-kata sendiri
2	3.4.6 membandingkan unsur kebahasaan beberapa teks prosedur lisan dan tulis dengan memberi dan meminta informasi terkait manual pendek dan sederhana, sesuai dengan konteks penggunaannya.	Disajikan sebuah teks prosedur tulis terkait **manual** pendek dan sederhana, peserta didik dapat menentukan makna kata.

Ingredients:

- 1 tb margarine, dairy free
- 1 tb Vegetable oil
- 1 c onion – chopped
- 8 oz Mushrooms – sliced
- 1 lg Garlic clove – minced
- 1/3 celery – thinly sliced
- 3 c Soybean milk
- 4 c Potatoes – cubed ½ inch
- ¼ ts Black pepper
- 1 ts Salt, or to taste
- 1 tb Parsley – chopped

Directions:

1. Pour onion, garlic, celery, and green pepper in hot oil and margarine for about 3 minutes. Add sliced mushrooms and sauté 3 more minutes.
2. Add soybean milk, black pepper, salt, and potatoes. Bring the mixture almost to a boil; reduce heat and simmer for 25 minutes, stirring occasionally.
3. Make about 6 cups. Thin with extra soybean milk if the soup is too thick.
4. Garnish with parsley and serve with the veal.

Level Cognitive C2 (LOTS)

1tb parsley – chopped. The word "chopped" is closest in meaning to …

A. Drawn

B. Hacked

C. Poured

D. Stirred

E. Cut

Level Cognitive C4 (HOTS)

The text states that …

A. ***extra soybean milk must be added if the soup is too thick***

B. six cups of soybean milk must be added to the soup

C. soybean milk is used merely for garnish

D. extra cups must be provided for everybody

E. The soup must be thick to make it easy

e. Formulating an interesting and contextual stimulus

The stimulus that is employed has to be interesting, which means that it has to be able to motivate students to read the stimuli. Stimuli that are interesting are typically new, have never been read by students, or are relevant to current issues.. Meanwhile, the contextual stimulus means a type of stimulus that corresponds to real life or daily life which can encourage the students to read it.

Several things need to be considered in preparing HOTS stimulus questions:

1. choose some information in the form of pictures, graphs, tables, discourses, etc which relate to the specific case;
2. the stimulus should fulfill the ability to interpret, and seek relationships,
3. analyze, conclude, or create;
4. choose the contextual and interesting (recent) cases/problems that motivate students to read;
5. stimulus must be directly related to the question (subject matter), and function
6. write question items in accordance with the question grid

Questions must be written in accordance with the criteria established in the question grid. The problems of HOTS are typically in the form of cases, so the first thing that must be considered is the use of appropriate stimulus and context. The process of selecting a matter should be considered based on the dimensions of the thought process that will be measured, such as the student's ability to analyse, evaluate, or create.

f. Determining the answer key or arranging rubric/scoring guidelines.

The key answer is the correct answer to the item test. For multiple-choice tests, matching tests or true and false that were answered correctly gave a score of 1 and 0 if the item was answered wrong / no answer. Arrange the scoring rubric (specific indicator/ specific description) for the subjective or essay test item. Following is the example of a scoring rubric for a writing test adapted from Brown 2007.

Components of Writing	Score	Level	Indicators	Weighting
Content (C)	4	Excellent	Present the information well chosen details across the paragraph	
	3	Good	Present the information with details in parts of the paragraph	3
	2	Fair	Present the information with some details	
	1	Poor	Present no clear information	
Vocabulary (V)	4	Excellent	Good in vocabulary choice	
	3	Good	Error in vocabulary choice are few and do not interfere with understanding	
	2	Fair	Error in vocabulary choice are and sometimes they interfere with understanding	2.5
	1	Poor	Many error in vocabulary choice that severally interfere with understanding	
Grammar (G)	4	Excellent	Good in grammar	
	3	Good	Error in grammar choice are few and do not interfere with understanding	
	2	Fair	Error in grammar choice are and sometimes they interfere with understanding	2.5
	1	Poor	Many error in grammar choice that severally interfere with understanding	
Mechanics (M)	4	Excellent	Good in spelling, punctuation and capitalization	
	3	Good	Error in spelling, punctuation and capitalization are few	
	2	Fair	Error in spelling, punctuation and capitalization, and sometimes interfere with understanding	2
	1	Poor	Error in spelling, punctuation and capitalization and severely interfere with understanding	

Final Score : $\frac{3C+2.5V+2.5G+2M}{40} x100$

Meanwhile, the scoring rubric of the speaking test can be constructed based on some indicators dealing with the

characteristics of speaking skills. One of the examples of scoring rubric of speaking can be adapted from O'Malley and Pierce (1990) as the following;

Aspects	Score	Criteria	Indicator
Pronuncia-tion and Intonation	1	Poor	Makes frequent problems with pronunciation and intonation
	2	Fair	Pronunciation and intonation are sometimes not clear or accurate
	3	Good	Pronunciation and intonation are clear or accurate
	4	Excellent	Pronunciation and intonation are almost always clear or accurate
Fluency	1	Poor	Hesitates too often when speaking, which interferes the communication
	2	Fair	Speaks with some hesitation, which often interferes the communication
	3	Good	Speaks with some hesitation, which seldom interferes the communication
	4	Excellent	Speaks smoothly, with little hesitation that does not interfere the communication
Accuracy	1	Poor	Uses basis structures and makes frequent errors
	2	Fair	Uses a variety of structures with frequent errors, or uses basic structures occasionally
	3	Good	Uses a variety of grammatical structures, but makes some errors
	4	Excellent	Uses a variety of structures with only occasional grammatical errors
Vocabulary	1	Poor	Uses only basic vocabulary and expressions
	2	Fair	Uses limited vocabulary and expressions
	3	Good	Uses a variety of vocabulary and expressions, but makes some errors in the word choice
	4	Excellent	Uses a variety of vocabulary and expressions and almost never makes some errors in the word choice

Source: https://www.researchgate.net/figure/Speaking-Rubric-Proposed-by-OMalley-and-Pierce-1990_tbl1_337066455

g. Writing an item card

Following are some types of item cards that can be used as a baseline in creating item cards.

Table 3.6 HOTS Item Card

Format subjects: English Subject

Class / Semester: XI/I

Basic competencies	: 3.5 menerapkan fungsi sosial, struktur teks, dan unsur kebahasaan teks interaksi transaksional lisan dan tulis yang melibatkan tindakan memberi dan meminta informasi terkait pengandaian diikuti oleh perintah/saran, sesuai dengan konteks penggunaannya. (Perhatikan unsur kebahasaan *if* dengan *imperative, can, should*)
Subject matter	: *Text* unsur kebahasaan *if* dengan *imperative, can, should)*
Indicator Problem	: Menentukan *gambaran umum* dari teks unsur kebahasaan *if* dengan *imperative, can, should)*
Cognitive level	: C4

Item Problem:

33. “If you want to pass the exam, you have to study harder.”If” clause used in the sentence is…

A. If clause + remainder

B. If clause + imperative

C. If clause + dream

D. If clause + suggestion E. If clause + general truth

Table 3.7 HOTS Item Card

<table>
<tr><td colspan="4">KARTU SOAL
ULANGAN TENGAH SEMESTER
Jenis Sekolah : SMA
Bahan Kelas / Semester: XI/I
Mata Pelajaran : Bahasa Inggris
Kurikulum : K13</td></tr>
<tr><td rowspan="2">Kompetensi Dasar
3.5 menerapkan fungsi sosial, struktur teks, dan unsur kebahasaan teks interaksi transaksional lisan dan tulis yang melibatkan tindakan memberi dan meminta informasi terkait pengandaian diikuti oleh perintah/saran, sesuai dengan konteks penggunaannya.
(Perhatikan unsur kebahasaan if dengan imperative, can, should)</td><td>No. soal : 33
Level Kognitif: C4</td><td>Kunci Jawaban: A</td><td>Produk / Proses: PG</td></tr>
<tr><td colspan="3">33. “If you want to pass the exam, you have to study harder.” If clause used in the sentence is…
a. If clause + remainder
b. If clause + imperative
c. If clause + dream
d. If clause + suggestion
e. If clause + general truth</td></tr>
<tr><td>Materi
Text unsur kebahasaan if dengan imperative, can, should)</td><td colspan="3"></td></tr>
</table>

Indikator Menentukan *gambaran umum* dari teks unsur kebahasaan *if* dengan *imperative, can, should)*	

Standard Competence	Item Number	Type of test	Cognitive Level
Peserta didik mampu mengidentifikasi konteks, gagasan utama, dan informasi terperinci dari ragam teks lisan/tulisan dalam bentuk Descriptive Text dalam lingkup keluarga dan kelas dengan benar.	19	MC	C5
Material: Descriptive Text	**Apriliana is my classmate. She is thirteen years old. She has a tall body. She is not fat. Her hair is black, straight, and long. She is very helpful to her friends. Her friends often**		

Indicator: Siswa dapat menyeleksi kalimat deskriptif yang sesuai untuk melengkapi sebuah descriptive text berupa deskripsi seseorang atau Binatang	**ask her for help to solve English problems. All her friends and teachers like her. She likes reading stories, cycling, gardening, and playing badminton.** **19.** Pay attention to the following statements. (1) The writer is twelve years old. (2) Apriliana is the writer's neighbor. (3) Apriliana is the writer's classmate. (4) Apriliana has a short and fat body. (5) Apriliana is friendly and really helpful. The **correct statements** based on the dialog are shown by number... a. (1) and (2) b. (2) and (3) c. (3) and (4) d. (3) and (5)

KARTU SOAL

Jenis Sekolah : SMA Negeri 3 Parepare
Mata Pelajaran : Bahasa Inggris
Bahan Kls/Smt : XII / 1 (Gazal)
Bentuk Soal : Pilihan Ganda
Aspek yang diukur : Pemahaman Text

Penyusun : 1. SUDIRMAN, S.Pd.
2. .....
Tahun Pelajaran : 2012/2013

KOMPETENSI DASAR :

5.2 Merespon makna dan langkah retorika dalam esei yang menggunakan ragam bahasa tulis secara akurat, lancar dan berterima dalam konteks kehidupan sehari-hari dan untuk mengakses ilmu pengetahuan dalam teks berbentuk: *narrative,*

MATERI :

The Monolog Text in the form of Narrative

INDIKATOR SOAL

49. Menemukan synonym kata
50. Menentukan tujuan komunikasi

BUKU SUMBER : Developing English Competencies for Senior High School, Grade XII of Natural and Social Science Program, Ahmad Doody, dkk. Pusat Perbukuan : Depdiknas 2008

RUMUSAN BUTIR SOAL

49. A group of frogs were travelling through *the woods*, and two of them fell into a deep pit. The synonym of the italicized word is … .

a. Sea
b. Trunk
c. Forest
d. Mount
e. Town

50. Which is the best lesson suited to the story?
a. Don't judge the book from its cover.
b. There is power of life and death in the tongue.
c. Be careful of what you say.
d. Early bird catches the worm.
e. A great talker is a great liar.

NO SOAL :	49	50
KUNCI :	c	b

KETERANGAN SOAL

NO	DIGUNAKAN UNTUK	TANGGAL	JUMLAH SISWA	TK	DP	PROPORSI PEMILIH						KET.
						A	B	C	D	E	OMT	
1.	Ulangan Akhir Semester											

Chapter 4
TOS (Table of Specifications)

A. What is TOS

TOS/ The table of specifications is a blueprint for a test. According to Mahrens and Lehman (1993), TOS is a blueprint of a test that involves the content area/material to be covered and the relative emphasis to be placed based on the area/ materials and instructional objectives stated in the curriculum. Gronlund and Linn (2000) state that a table of specifications is a material of course stated in a certain curriculum that can be broadly set to put into both a subject matter content and instructional objectives. A table of specifications is also can be defined as a kind of chart used by the teacher in writing item tests. It will guarantee the test really evaluates the material taught and the learning experience allowed to the students. This situation also straightens the instrument of the test with the learning objectives and the cognitive level will be measured. In addition, it also describes the number of item test which assesses every learning objective. In short, the table of specifications defines the area and the topics of the test (Generator, G: 2020). Thus, a table of specifications (TOS) is a kind of tool used to determine whether a test or assessment really measures the content and the student's thinking skills or not. The great test

intends to measure what really has to be measured. In short, TOS is a plan prepared by the teacher as a basis to construct the test.

There are some purposes for making a table of specifications. They involve:

(1) to identify the cognitive domain that will be measured and to ensure that a representative and adequate sample of questions arises on the test.
(2) to provide a blueprint to ensure that the teacher in this case has a balanced measure of their course learning objective.
(3) permit the teacher to compose a test that concentrates on the key areas/materials and weights those different areas/materials based on their significance.

Generate. G (2020) describes in detail the purpose of making a table of specifications is to permit the teacher to;

a. set the content area or material of test
b. construct the item test that portrays the students' learning over the cognitive level
c. allocate a specific amount of instructional time needed for every topic/SLO
d. confirm that none of the learning objectives left behind to be measured
e. construct valid item test
f. evaluate learning objectives that actually belong to instructional process

g. determine proportional emphasis on each learning objective

Fives, Helen rose and DiDonato-Barnes, Nicole (2013) stated the goal of this strategy: improving the validity of a teacher's evaluations based on a given assessment.

Based on the description above, it can be concluded that TOS is a tool to help the teacher decide to create test construction and improve the validity of the teacher-made test for evaluation. Content validity is the main output that directly can be obtained from TOS. It also helps the teacher to identify the types of tests they need based on the course of materials.

B. The Component of TOS

A table of specifications /TOS can be used by the teachers to help them plan and write their test construction. A table of specifications is a kind of guidance to assist a teacher or examiner in conducting an evaluation (Akem: 2006). It is very useful for the teacher because it's constructed based on some criteria. Those are;

a. learning objective
b. topics of materials discussed in class
c. time spent discussing the topics
d. topics provided in the textbook chapter
e. the interest and capacity put in the text

The TOS provides a framework for organizing information about the instructional activities experienced by the student (Fives, Helen Rose and DiDonato-Barnes, Nicole:2013) Before constructing the TOS, teachers need to determine (1) the number of test items to include and (2) the distribution of multiple-choice and short-answer items. Limiting the levels of cognitive processing to high and low levels, rather than dividing out across the six levels of cognitive processing described by Bloom (1956) and updated by Anderson et al (2001), is one way to simplify the TOS. This can be done by limiting the levels of cognitive processing to high and low levels.

C. How to Construct TOS

The TOS is arranged as a table containing some key information to help the teachers straighten the learning objectives that represent the content or material of learning and the cognitive levels proposed for the students to achieve the learning outcomes with the class time spent and the number of test items.

There are some types of tables of specifications that can be created by the teacher before writing the item test.

(1) The first type of TOS was adapted from the Generate style that was proposed in 2020. He recommended a way to construct a table of specifications as the following steps;

a. **Tabulate the topic, SLOs, and cognitive level of the taught curriculum**.

Table 4.1: Table specification of Mid-term Test

TOPICS	SLOs	ANALYZE	EVALUATE	CREATE
Descriptive Text	5	2	2	1
Announcement Text	5	2	1	2
Procedure Text	6	4	1	1
Narrative Text	4	3	1	-
Total number of SLOs	**20**	**11**	**5**	**4**
Percentage	100	55	25	20
Mark	40	22	10	8
Total Mark	80			

The table above describes the specifications of the mid-term test instrument. The first column represents the material taught by the teacher, the second column shows the students' learning outcomes / SLOs during those periods, and the last three columns represent the cognitive level. For the table above, the

cognitive level focuses on the highest-order thinking skill (Analyse, Evaluate, and Create).

b. **Tabulate the percentage share of each topic and the distribution of marks for each topic**.

After finishing the first step, then prepare for the next step for Table 2. It represents the topics or material taught by the teacher, no. of SLOs for each topic, the percentage share of SLOs, and the distribution of marks. The following formulas can be used to calculate the percentage share of SLOs and the distribution of marks.

Weight in percentage = No of SLOs of the topic x 100: total number of SLOs

Weight in Marks = Calculated % of max marks

Table 4.2: content distribution of test syllabus

TOPICS	No SLOs in a topic	% share of SLOs	Distribution of Marks
Descriptive Text	5	25 %	20
Announcement Text	5	25 %	20
Procedure Text	6	30 %	24

Narrative Text	4	20 %	16
TOTAL	20	100 %	80

c. Tabulate item distribution per cognitive level takes place.

Calculating item distribution for cognitive levels can be done by using the following formula;

The number of items = max marks topics: no. of SLOs x number of SLOs in each cognitive level

In distributing item tests at the cognitive level, teachers have to really consider the characteristics of each level. There are two types of models that can be used by the teacher as a baseline in distributing the items.

Generate. G (2020) has a different point of view on the following case described in the table below;

Table 4.3. Item Distribution Per Cognitive Level (Generate. G version)

TOPICS	Allocated Marks	MCQs	K	U	A
Descriptive Text	20	5	2	2	1
Announcement Text	20	5	2	1	2
Procedure Text	24	6	4	1	1
Narrative Text	16	4	3	1	-
	80	20	**11**	**5**	**4**

All cognitive levels of Bloom's taxonomy are not covered in the table above. On the contrary, the teachers of secondary level and higher levels distribute the item in three levels, they involve knowledge, understanding, and application. In this case, they suppose all higher-order thinking skills as a part of an application. The next model of item distribution on a cognitive level is distributed based on the HOTS version. This distribution gives emphasis on higher-order thinking skills which involve the level of analysis, evaluation, and creation as described in the following table:

Table 4.5 Item Distribution Per Cognitive Level (HOTS version)

TOPICS	Allocated Marks	Essay	MCQs	Analysis	Evaluation	Creation
Descriptive Text	20	1	4	2	2	1
Announcement Text	20	2	3	2	1	2
Procedure Text	24	1	5	4	1	1
Narrative Text	16	-	4	3	1	-
	80	4	16	**11**	**5**	**4**

The table above describes in detail the component of the test. The table of specifications above determines the quality and credibility of the test.

(2) Table of specifications is constructed based on the learning objective and the contents of the course have been established by the teacher. Miller et al. (2012) proposed a simple way of creating a table of specifications. Through three steps a table of specifications will be available. It involves;

a. List the general learning objectives across the top of the table.
b. List the high-level content areas down the left side of the table.
c. Determine what proportion of the test items should be devoted to each objective and each content area.

(3) According to Esquerra.J. Rovin (2003) steps that must be done by the teachers in composing TOS. They have to allow the following steps;

1. make a list of the topic that will be used in a test
2. establish the learning objectives that will be assessed
3. decide the number of days/hours spent teaching a specific topic
4. establish the percentage allocation of the test items for each of the topics covered by using the following formula;

$$\%\ of\ topic = \frac{Total\ number\ of\ hours\ spent}{Total\ number\ of\ hours\ spent\ teaching\ the\ topic}$$

5. set the number of items for every topic by multiplying the percentage allocation for every topic by the total number of items to be constructed.

6. Distribute the number of items into the objective learning based on the degree of importance attached by the teacher.

An example of finding the number of items for every topic can be done as the following;

The number of items for the topic of "Descriptive Text"

Number of class sessions: 2

Desired number of items: 25

Total number of class sessions: 10

$$N = \frac{\text{number of class sessions x desired number of items}}{\text{total number of classes}}$$

$$= \frac{2 \times 25}{10}$$

$$= 5$$

N = Number of items

So, the number of items for the topic of "descriptive text" is 5 item number.

Making TOS before writing the instrument of the test will make the instrument more systematic and organized. This process also will improve the validity of teachers' evaluations based on the constructed instrument used in their classroom evaluation process.

Without TOS, the instrument of the test created by the teacher will not meet the criteria of content validity. A test for this matter carries out content validity only if the tool or the instrument of the assessment really assesses what it is supposed to be measured or assessed.

D. Types of TOS (Table of Specification)

1. Types of TOS proposed by (Pearson Test of English General)

Test structure

Written test

Section	Skills	Item Types	Objectives						Score points
			Level A1	Level 1	Level 2	Level 3	Level 4	Level 5	
1	Listening	3-option multiple choice (A1 and L1 graphical options)	To assess ability to understand the gist of short spoken utterances			To assess ability to understand the main detail in short spoken utterances			10
2	Listening and Writing	Dictation	To assess ability to understand a short utterance by transcribing a spoken text		To assess ability to understand an extended utterance by transcribing a spoken text				5 listening 5 writing
3	Listening	Text, note completion	To assess ability to extract specific information from spoken texts		To assess ability to extract specific information from extended spoken texts				10
4	Reading	Gap fill 3-option multiple choice	To assess ability to extract specific information from spoken texts	To assess ability to understand the purpose, structure and main idea of short written texts					5
5	Reading	3-option multiple choice (A1 and L1 graphical options)	To assess ability to understand the main detail in short written texts		To assess ability to understand the main ideas in an extended written text				5
6	Reading	Open-ended question	To assess ability to understand the main points of short written texts			To assess ability to understand the main points of short and extended written texts			8
7	Reading	Text, note completion	To assess ability to extract specific information from a written text		To assess ability to extract specific information from an extended written text				7
8	Writing	Write correspondence	To assess ability to write a short piece of correspondence (30-50 words)	To assess ability to write a short piece of correspondence (50–70 words)	To assess ability to write a piece of correspondence (70–90 words)	To assess ability to write a piece of correspondence (90–120 words)	To assess ability to write a piece of correspondence (120–150 words)	To assess ability to write a piece of correspondence (150–200 words)	10
9	Writing	Write text	To assess ability to write a short text based on a picture (50-80 words)	To assess ability to write a short text based on a series of three pictures (80–100 words)	To assess ability to write a short text from own experience, knowledge or imagination (100–150 words)	To assess ability to write a short text from own experience, knowledge or imagination (150–200 words)	To assess ability to write a short text from own experience, knowledge or imagination (200–250 words)	To assess ability to write a short text from own experience, knowledge or imagination (250-300 words)	10

Speaking Test

Section	Skills	Item Types	Objectives						Score points
			Level A1	Level 1	Level 2	Level 3	Level 4	Level 5	
10	Speaking	Discussion	Not featured at this level		To assess ability to discuss a concrete issue (2 mins)		To assess ability to discuss a concrete or abstract issue (2 mins)		25 score points in total; distributed across the marking criteria
11	Speaking	Describe picture	To assess ability to speak about a picture (2 mins)		To assess ability to speak continuously about a picture and interpret some aspect of it (1.5 mins)	To assess ability to speak continuously about two related pictures and interpret some aspect of them (1.5 mins)			
12	Speaking	Role play	To assess ability to perform and respond to basic language functions appropriately (1.5 mins)		To assess ability to perform and respond to language functions appropriately (1.5 mins)	To assess ability to perform and respond to language functions appropriately (2 mins)	To assess ability to speak continuously on a topic related to two pictures (2 mins)		

The following are the scoring rubrics for the TOS above;

that can be obtained in the written component of the test at all levels.

Sections	Skills	Score points
1	Listening	10
2	Listening and writing	10 (5 listening, 5 writing)
3	Listening	10
4	Reading	5
5	Reading	5
6	Reading	8
7	Reading	7
8	Writing	10
9	Writing	10
Total		**75**

Sections	Skills	Score points
10	Speaking (only featured at levels 2-5, NOT levels A1 and 1)	**25 score points** in total; distributed across the marking criteria
11	Speaking	
12	Speaking	
Total		**25**

2. Types of TOS proposed by Nortar

It provides a two-way chart that has the purpose of helping the teachers relate their instructional objectives, the cognitive level of instruction, and the number of tests that should assess each objective (Nortar et al., 2004).

Table 3: A Sample Table of Specifications for Fifth Grade Social Studies Chapter 6: The Middle Colonies

A	B	C	D	E	F	G
	Instructional Objectives	Time Spent on Topic (minutes)	Percent of Class Time on Topic	Number of Test Items: 10	Lower Levels -Knowledge -Recall -Identification -Comprehension	Higher Levels -Application -Analysis -Evaluation -Synthesis
Day 1	1. Identify the various groups who settled the Middle Atlantic Colonies.	15	10.00%	1.00	1 Multiple Choice	
	2. Summarize the contributions of different religious and cultural groups to the settlement of the Middle Atlantic Colonies.	15	10.00%	1.00	1 Short Answer	
Day 2	3. Identify George Whitefield as an early leader of the Great Awakening	10	6.70%	.67	1 Multiple Choice	
	4. Evaluate the impact of the Great Awakening sermons on English colonists.	20	13.30%	1.33		1 Multiple Choice
Day 3	5. Describe the physical features that helped Philadelphia become a main port.	15	10.0%	1.00	1 Multiple Choice	
	6. List ways in which immigrants aided Philadelphia's growth and prosperity.	10	6.70%	.67	1 Short Answer	
	7. Identify the contributions Benjamin Franklin made to Philadelphia.	5	3.30%	.33	---	
Day 4	8. Interpret information in a circle graph.	15	10.00%	1.00		1 Multiple Choice
	9. Gather and organize information using a circle graph.	15	10.00%	1.00		1 Short Answer
Day 5	10. Identify the challenges faced by backcountry settlers.	5	3.30%	.33	---	
	11. Analyze the importance of the Great Wagon Road as an early transportation route.	10	6.70%	.67		1 Multiple Choice
	12. Explain how backcountry settlers adapted to and made use of the resources available to them.	15	10.00%	1.00		1 Multiple Choice
		150	**100.00%**	**10**	**5**	**5**

3. Types of TOS constructed based on Indonesia Curriculum

No	ATP / Basic Competence	Material	Class	Indicator	Cognitive Level	Item Type	Item Number	Item score
1.	Applying text structures and linguistic elements correctly from various texts regarding asking simple questions in the form of greeting & parting, introducing oneself and others within the family and class.	Expression about greeting and parting Expression about introducing oneself & others	VII	1. Determine greetings and parting that are appropriate to the context of everyday life. 2. Choose the correct introduction sentence to complete the dialogue and paragraph 3. Develop the correct order of asking and giving questions in greeting and introduction sentences.	C-4 (Analyze) C-5 (Evaluate) C-6 (Create)	Multiple choice Multiple Choice Essay	1,2,5,6 3, 4, 7,8	2 2 rubric
2.	Students are able to identify the context, main idea, and detailed information from a variety of spoken/written texts in the form of Descriptive Text within the family and class correctly.	Descriptive text	VII	1. connecting incomplete descriptive sentences in a descriptive text. 2. Select appropriate descriptive sentences to complete a descriptive text in the form of a description of a person or animal. 3. selecting appropriate descriptive sentences to complete a descriptive text in the form of a description of a person or animal 4. Rearrange the descriptive text that has been presented randomly into one text that matches the characteristics of the descriptive text.	C-4 (Analysis) C-5 (Evaluate) C-5 (Evaluate) C-6 (Create)	Multiple Choice Multiple Choice Multiple choice Essay	9,10,11, 12,13 14,15,16,17 18,19,20 30	2 2 2 rubric

References

Anderson dan Krathwohl. 2001. A Taxonomy for Learning, Teaching, and Assessing (A Revision of Bloom's Taxonomy of Educational Objectives). Abridge Edition. Penerbit David McKay Company. New York

Anderson, L.W., Krathwohl, D.R., Airasian, P.W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., Raths, J., Wittrock, M.C. (2001). A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. New York: Pearson, Allyn & Bacon

Bloom, B.S. (Ed.). Engelhart, M.D., Furst, E.J., Hill, W.H., Krathwohl, D.R. (1956). Taxonomy of Educational Objectives, Handbook I: *The Cognitive Domain.* New York: David McKay Co Inc.

Brookhart, S.M. (2010) How to Assess Higher-Order Thinking Skills in Your Classroom. Alexandria, VA: ASCD.

Brown, H.D. (1994). Principles of Language Learning and Teaching. New Jersey: Prentice-Hall.

Cannon, H. M. and Feinstein, A. H (2005). Bloom Beyond Bloom: Using the Revised Taxonomy to Develop Experiential Learning Strategies, Developments in Business Simulation and Experiential Learning, Vol. 32, 2005

Content Generator. Table of Specifications, Definition, development, and Implementation. WWW. CONTENTGENERATOR.COM. December 12, 2020. July 18, 2023.

Cox, Janelle (2019). Teaching Strategies that Enhance Higher-Order Thinking. K-12 Resources By Teachers, For Teachers.https://www.teachhub.com/teachingstrategies/2019/10/teaching-strategies-that-enhance-higher-order-thinking/

Doff, A. (1988). Teach English. Glasgow: Cambridge University Press.

Esquerra, Rovin. J. Preparing Table of Specifications. https://www.studocu.com/ph/document/the-college-of maasin /bachelor-of-secondary-education/preparing-table-of-specifications/26496224 Accessed on July 28, 2023.

Fives, Helenrose and DiDonato-Barnes, Nicole (2013) "Classroom Test Construction: The Power of a Table of Specifications," Practical Assessment, Research, and Evaluation: Vol. 18 , Article 3. DOI: https://doi.org/10.7275/cztt-7109 Available at: https://scholarworks.umass.edu/pare/vol18/iss1/3

Ganapathy, M., & Kaur, S. (2014). ESL Students' Perceptions of the use of Higher Order Thinking Skills in English Language Writing. Advances in Language and Literary Studies, 5(5), 80–87. https://doi.org/10.7575/aiac.alls.v.5n.5p.80

Heong, Y. M.,Othman, W.D.,Md 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, Vol. 1,No. 2, July 2011, 121- 125

Krathwohl, D.R., Bloom, B.S., Masia, B.B. (1973). Taxonomy of Educational Objectives, the Classification of Educational Goals. Handbook II: Affective Domain. New York: David McKay Co., Inc.

Kuswana, W. S. (*2012*). Taksonomi Kognitif Perkembangan Ragam Berpikir. Bandung: PT Remaja Rosdakarya Offset.

Le Grange, L.L. & Reddy, C. (1998). Continuous Assessment: An Introduction and Guidelines to Implementation. Cape Town, South Africa: Juta

Layosa, Roxette U. 2014. Preparing table of Specification. https://www.slideshare.net/roxet02/preparing-a-tos. Accessed on July 21, 2023.

Mehrens, *W.A., &* Lehmann, *I.J.* (1973). Measurement and evaluation in education and psychology. New York: Holt, Rinehart, and Winston. Inc.

Miller, M.D., Linn, R.L. and Gronlund, N.E. (2000) Measurement and Assessment in Teaching. 10th Edition, Pearson Education Ltd., Upper Saddle River.

Miller, M. D., Linn, R. L., & Gronlund, N. E. (2012). Measurement and assessment in teaching (11th ed.). Upper Saddle River, NJ: Pearson Education, Inc.

Nur, Ikhwan Rois (2019). HIGHER ORDER THINKING SKILL (HOTS) ASSESSMENTS ON LEARNING ENGLISH. Proceeding of 1 st Conference of English Language and Literature (CELL)

Notar, C.E., Zuelke, D. C., Wilson, J. D. & Yunker, B. D. (2004). The table of specifications: Insuring accountability in teacher made tests. Journal of Instructional Psychology, 31, 115-129.

Rosset, E. (1998). Fill in the Gaps. Brno: Didaktis.

Sani, Ridwan Abdullah. (2019). Pembelajaran Berbasis HOTS (Higher Order Thinking Skills). Tangerang: Tira Smart.

Schraw, Gregory, Robinson, DH. 2011. Assessment Of Higer Order Thinking Skillss. America: Information Age Publishing.

Scrivener, J. (1998). Learning Teaching. Oxford: Macmillan.

Sincero, S. M (April 18, 2011). Domains of Learning. Accessed from https://explorable.com/domains-of-learning Date accessed 8th October 2018.

Susanti, et al. (2021). Panduan Pembelajaran dan Asesmen Jenjang Pendidikan Dasar dan Menengah (SD/MI, SMP/MTs, SMA/SMK/MA) Pusat Asesmen dan Pembelajaran Badan Penelitian dan Pengembangan dan Perbukuan Kementerian Pendidikan Kebudayaan, Riset, dan Teknologi

Sousa, D. A (2016). How the Brain Works. Crowin Press. 2016.

Ur, P. (1991). A Course in Language Teaching. Cambridge: Cambridge University Press.

Tenrycolle. 2022. Education Testing: Definition, Types & Task. Tenrycolle.com.https://tenrycolle.com/materials/universit

as/education-testing-definition-types-task/. Accessed on July 23, 2023.
The University of Kansas. Table of Specification. www. specialconnections.ku.edu. Department of Special Education. Accessed on August 2, 2023.
Weiss, C.H. (1972). Evaluation Research: Methods of Assessing Program Effectiveness. Englewood Cliffs (NJ), USA: Prentice-Hall.
Widana, I Wayan. (2016). Penulisan Soal HOTS untuk Ujian Sekolah. Jakarta: Direktorat Pembinaan SMA

Authors' biography

Nuri Ati Ningsih was born in Madiun on February 14, 1975. She graduated from English Language Education Department, at Muhammadiyah Malang University in 1997. She continued her master degree in Educational Technology at Adibuana University and graduated in the year of 2007. Then, in 2013 she also got her master degree in English Education from Sebelas Maret University Surakarta. Her career as a lecturer at the English Education Department of UNIPMA was started in 1999 up to now.

She shares her knowledge and experience through writing some books, such as *Evaluation on Language, Language Testing 1, Language Testing 2, Boundless English, and Community Language Learning,Understanding Reading Text in English for University Student, and Kesantunan Berbahasa Perempuan di Jejaring Sosial.*

Rengganis Siwi Amumpuni was born in Madiun on August 14, 1984. She graduated from English Teaching Department, at IKIP PGRI Madiun in 2008. She earned her Master degree on English Education at Universitas Sebelas Maret Surakarta in 2012. She is now a doctoral candidate at Universitas Negeri Semarang, Indonesia. She is an awardee of Beasiswa Pendidikan Indonesia batch 2021 by Balai Pembiayaan Pendidikan Tinggi (BPPT) and Lembaga Pengelola Dana Pendidikan (LPDP). Her career as a lecturer at the English Education Department of UNIPMA was started in 2010 up to now.

She shares her knowledge and experience through writing some books, such as *Demystifiying Language Teaching Program Management and Evaluation, Intensive Reading, and Basic Listening*, *and Mengenal Bahasa Inggris melalui Benda Sekitar.*

This book is designed to be used as a reference for teachers with skills and knowledge about constructing a table of specification (TOS) as a blueprint for a test. TOS is a tool to help the teacher decide to create test construction and improve the validity of the teacher-made test for evaluation. It is also helpful for teachers to determine the kinds of examinations they need to administer based on the course materials. It involves the content material to be covered and the relative emphasis to be placed based on the materials and instructional objectives stated in the curriculum. The contents of this book were taken from various relevant and up-to-date data source which covers some chapter including testing, evaluating, and assessment in language learning, bloom taxonomy, testing by using HOTS (higher order thinking skills) and constructing test specification.