AN ANALYSIS ON CONTENT VALIDITY OF ENGLISH ACHIEVEMENT TEST BASED ON THE SYLLABUS FOR THE FIRST GRADE OF MADRASAH TSANAWIYAH ANNAJAH PETUKANGAN

A “Skripsi”
Presented to the Faculty of Tarbiya and Teachers’ Training in a Partial Fulfillment of the Requirements for the Degree of Strata I (S1) in English Language Education

By:
NURLIA GUSTIARI
NIM: 105014000315

DEPARTMENT OF ENGLISH EDUCATION
THE FACULTY OF TARBIYA AND TEACHERS’ TRAINING
“SYARIF HIDAYATULLAH” STATE ISLAMIC UNIVERSITY
JAKARTA
2010
AN ANALYSIS ON CONTENT VALIDITY OF ENGLISH ACHIEVEMENT TEST BASED ON THE SYLLABUS FOR THE FIRST GRADE OF MADRASAH TSANAWIYAH ANNAJAH PETUKANGAN

A “Skripsi”

Presented to the Faculty of Tarbiya and Teachers’ Training in a Partial Fulfillment of the Requirements for the Degree of Strata I (S1) in English Language Education

By:
NURLIA GUSTIARI
NIM: 105014000315

DEPARTMENT OF ENGLISH EDUCATION
THE FACULTY OF TARBIYA AND TEACHERS’ TRAINING
“SYARIF HIDAYATULLAH” STATE ISLAMIC UNIVERSITY
JAKARTA
2010
ENDORSEMENT BY THE EXAMINATION COMMITTEE

The Paper entitled “An Analysis on Content Validity of English Achievement Test Based on the Syllabus for the First Grade of Madrasah Tsanawiyah Annajah Petukangan”, written by Nurlia Gustiari, student’s registration number 105014000315 was examined in the examination session of the Faculty of Tarbiya and Teachers’ Training, “Syarif Hidayatullah” State Islamic University Jakarta on February 24th 2010. The Paper has been accepted and declared to have fulfilled one of the requirements for the degree of S.Pd. (S1) in the Department of English Education.

Jakarta, March 4th 2010

Examination Committee

Chairman: Drs. Syauki, M.Pd.
NIP. 196412121991031002

Secretary: Neneng Sunengsih, S.Pd.
NIP. 197306251999032001

Examiners:
1. Drs. H. AM. Zaenuri, M.Pd.
NIP. 195303041979031003

2. Drs. H. Bahrul Hasibuan, M.Ed.

Acknowledgement by:
Dean of the Faculty of Tarbiya and Teachers’ Training

Prof. Dr. Dede Rosyada, MA.
NIP. 195710051987031003
AN ANALYSIS ON CONTENT VALIDITY OF ENGLISH ACHIEVEMENT TEST BASED ON THE SYLLABUS FOR THE FIRST GRADE OF MADRASAH TSANAWIYAH ANNAJAH PETUKANGAN, THE 2008/2009 ACADEMIC YEAR

A “Skripsi”

Presented to the Faculty of Tarbiya and Teachers’ Training in a Partial Fulfillment of the Requirements for the Degree of Strata 1 (S1) in English Language Education

Approved by the Advisor

Drs. Syauki, M.Pd.
NIP. 196412121991031002

DEPARTMENT OF ENGLISH EDUCATION
THE FACULTY OF TARBIYA AND TEACHERS’ TRAINING
“SYARIF HIDAYATULLAH” STATE ISLAMIC UNIVERSITY
JAKARTA
2010
ABSTRACT


Pembimbing: Drs. Syauki, M.Pd.

**Key Words:** Summative Test, Validity, MTs Annajah

Evaluation plays an important role in teaching learning activities. It is an integral part of the instructions program. Through evaluation, teachers are able to find out the effectiveness or the failure of a method and also students achievement in mastering the lesson.

This research is to find empirical evidence of the English summative test validity made by the English teacher of MTs Annajah Petukangan for even semester of the first grade students. The total number of the test items is 45 items which consist of 40 multiple choice and 5 essay test items. This research used description and evaluation analysis.

The findings of the research prove that the items of English summative test for even semester of the first grade students in MTs Annajah have good content validity. It shows that the materials of the English summative test are appropriate to the curriculum which is recommended.

Based on the findings of the research, it can be suggested that (1) the test makers should have a good preparation before making a test, by considering the principles of constructing good test items, (2) the test makers should consider that the test items should be appropriate to the syllabus category used, the items that do not fulfill to the curriculum should be discarded, (3) While the items that do not fulfill to the syllabus should be revised so that they can be used for the next evaluation.
ACKNOWLEDGEMENT

In the name of Allah, The Beneficent, The Merciful

All praise be to Allah, the Lord of The World, The Almighty God for blessing, guidance, help and love so the writer can complete this “skripsi” properly. Peace and salutation for our prophet Muhammad, Peace be Upon Him, and to his family, relatives as well as his companion.

This “skripsi” is presented to English Education Department, Faculty of Tarbiya and Teachers’ Training, “Syarif Hidayatullah” State Islamic University Jakarta, as a partial fulfillment of the requirements for the degree of strata 1 (S1). In this “skripsi” the writer analyzes the content validity of English summative test items at even semester of the first grade students at Madrasah Tsanawiyah Annajah Petukangan, the 2008/2009 academic year.

In finishing this “skripsi”, the writer got many guidance and motivation from her advisor, Drs. Syauki, M. Pd., for his patience, valuable guidance, encouragement, comment and suggestions for completing her “skripsi”.

The writer would like to express her gratitude to:

1. Prof. Dr. Dede Rosyada M.A., the Dean of Faculty of Tarbiya and Teachers’ Training, his concern and spending of his best time to the special attention for better education regarding with FITK. It is one of tremendous aspects for best and conducive atmosphere at FITK.

2. Drs. Syauki, M. Pd., the Head of English Education Department, his concern and mind for better English education will be a historical point and considerable sample for the next terms.

3. Neneng Sunengsih, S. Pd., the Secretary of English Education Department, her best time and attention for English Education is so inspiring.

4. All the English Education Department Lecturers who had transferred their knowledge and also for their valuable guidance and encouragement to the writer.
5. Drs. Sam’unal Ghozi, the Principal of Madrasah Tsanawiyah Annajah Petukangan, to the special attention for her “skripsi”.

6. The English Teacher of Madrasah Tsanawiyah Annajah Petukangan, Ilfa Rianti, S.Pd., who is very patience to help her in doing her “skripsi”. Also the staff officer of Madrasah Tsanawiyah Annajah Petukangan, Maudi, who is very kind to help her in doing her paper.

7. All her beloved big family, especially her parents: her mother, Nurhayati, and her father, Bahrudin, for their pray, patience, their love and their support to the writer in her life. The writer also wishes to express her gratitude to her siblings, her beloved brothers, Fachrul Rhozi and Syaiful Anwar, who give support and permission to use the computer so long to the writer.

8. Thanks to her special someone, Bobby Aji Perdana, who always prays, supports, be patience, helps, and accompanies her in finishing the “skripsi”.

9. All her friends in “A class” of English Education Department 2005, which is very fantastic class. Especially for her best friends, Suci, Ena, Ika, Ulfa, Adaw, Reny, Qadarna and Dinda who help her in doing her “skripsi”.

Finally, the writer realizes that this “skripsi” is far from being perfect. Therefore, it is really a pleasure for her to suggest and critic from everyone who will encourage her to complete this “skripsi”. The writer also expects this “skripsi” will give valuable information for the development of scientific education and become the inspiration for people who read it.

Jakarta, 2 February 2010

The writer
TABLE OF CONTENTS

ABSTRACT .........................................................................................................................i
ACKNOWLEDGEMENT ................................................................. ii
TABLE OF CONTENTS ........................................................................ iv
LIST OF TABLES ........................................................................ vi
LIST OF PICTURE ........................................................................ vi
LIST OF APPENDICES ................................................................. vi

CHAPTER I INTRODUCTION
  A. The Background of the Study ......................................................... 1
  B. The Limitation of the Problem ...................................................... 4
  C. The Formulation of the Problem .................................................... 4
  D. The Objective of the Study ............................................................ 4
  E. The Method of the Study ............................................................... 5
  F. The Significance of the Study .......................................................... 5

CHAPTER II THEORETICAL FRAMEWORK
  A. Test ......................................................................................... 6
     1. The Meaning of Test ............................................................... 6
     2. Purposes of Test .................................................................... 7
     3. Types of Test and Types of Test Item ..................................... 8
     4. The Characteristics of a Good Test ........................................ 12
  B. Validity .................................................................................... 17
     1. Nature of Validity ................................................................. 17
     2. Types of Validity .................................................................. 18
  C. Achievement Test ...................................................................... 22
     1. The Construction of Achievement Test ................................. 22
     2. Basic Principles of Achievement Testing ............................... 22
     3. Types of Achievement Test ............................................... 23
D. Curriculum .................................................................................... 25
   1. Meaning of Curriculum........................................................... 25
   2. Purpose of Education............................................................ 26
   3. Approach of Curriculum......................................................... 27
E. English Curriculum........................................................................ 28
   1. Background.............................................................................. 29
   2. Competence Model ............................................................... 30
   3. Purpose.................................................................................... 31
   4. Scope ...................................................................................... 31
   5. Competence Standard and Basic Competency ....................... 32

CHAPTER III IMPLEMENTATION OF THE RESEARCH
A. The Methodology of the Research ................................................ 34
   1. The Objective of the Research ................................................ 34
   2. The Time and Location ........................................................... 34
   3. The Techniques of Data Collecting ........................................ 35
   4. The Techniques of Data Analysis ............................................ 35
B. The Findings of the Research and The Discussions ................... 36
   1. The Description of the Data .................................................... 36
   2. The Analysis of the Data ....................................................... 40
   3. The Interpretation of the Data ................................................. 56

CHAPTER IV CONCLUSIONS AND SUGGESTIONS
A. Conclusions ................................................................................ 59
B. Suggestions ................................................................................ 60

REFERENCES ................................................................................... 61
APPENDICES....................................................................................... 64
LIST OF TABLES

Table 2.1 The ‘Multi-Level’ or ‘Variable Focus’ Curriculum ...................... 28
Table 2.2 Competence Standard and Basic Competency ............................ 32
Table 3.1 Criteria of Test Result Percentage ........................................... 36
Table 3.2 The Appropriateness of English Summative Test with the
 indicators in the latest English syllabus at Even Semester of the
 First Grade Students at MTs Annajah ............................................ 37
Table 3.3 The Inappropriateness of English Summative Test with the
 indicators in the latest English syllabus at Even Semester of the
 First Grade Students at MTs Annajah ............................................ 39
Table 3.4 Question and Explanation Items Based on “SK”: Reading .......... 40
Table 3.5 Question and Explanation Items Based on “SK”: Writing .......... 46
Table 3.6 The Analysis Result of Appropriateness and Inappropriateness
 of English Summative Test Items at Even Semester of the First
 Grade Students of MTs. Annajah ................................................... 56
LIST OF PICTURE

Picture 2.1 Communicative Competence Model (by Celce-Murcia et al) ....... 30

LIST OF APPENDICES

Appendix 1  Question Sheets ................................................................. 64
Appendix 2  Blue Print of the Questions .............................................. 68
Appendix 3  English Syllabus ................................................................. 72
Appendix 4  Result of Interview ............................................................. 87
Appendix 5  Letters ............................................................................... 89
CHAPTER I
INTRODUCTION

This chapter presents and discusses the background of the study, the limitation of the problem, the formulation of the problem, the objective of the study, the method of the study, and the significance of the study.

A. The Background of the Study

Many people in the world speak English as a first language or as a second language. Formally, English is studied from Junior High School till University. In Indonesia, English subject especially in Junior High School has already become as a compulsory subject and it is one of the requirements to pass the national exam.

English as one of the main subjects in Junior High School has function as a developing instrument for the students in knowledge, technology and culture. So, they are ready to take a part in national development.

The purpose of teaching English in Indonesia is to develop the communication skills especially in oral and written skills (listening, speaking, reading, and writing). To measure the student’s skill in the teaching-learning process, the teachers need to hold an evaluation.

Evaluation is a very important aspect in teaching and learning activities. Evaluation plays an important role in some activities, especially in terms of education. Through evaluation a teacher will be able to know his or her students’ achievement on the materials that have been taught in a certain period of time, so the teacher can measure his or her effectiveness in teaching which has been applied in the classroom. And the information gained through the evaluation will be very useful to make improvement in the future.

According to Norman, “Evaluation may be defined as the systematic process of collecting, analyzing, and interpreting information to determine the extent to which pupils are achieving instructional objectives.”

---

1 Norman E. Gronlund, and Robert L. Linn, Measurement and Evaluation in Teaching, (New York:
One of the ways to do the evaluation is through a test. The test can measure the students’ ability. Test is a short examination of knowledge or ability, consisting of question that must be answered or activities that must be carried out.\(^2\)

To get a good result, the test which fulfills the criterion of a good test: that is validity, reliability and usability must be arranged well.\(^3\) In this time, the writer just focuses on the content validity because the writer wants to compare the test with the syllabus.

Based on the curriculum 2006, the objectives of English lesson in Junior High School are that the students have abilities like:

1. To develop communicating competence orally and written to reach functional literacy.
2. To have an awareness of English essence and important to enhance the national potency in global society.
3. To develop the comprehension of students about the relevancy between language and culture.\(^4\)

To reach the objectives in the English curriculum, the writer considers that the test is one of the instruments, which can be used for gathering the information about the strength and the weakness in accepting the lesson of the students. Teachers are ones who know the characteristics of their classes. Thus, they are in the best position to construct the test items to measure their students’ achievement and it is not an easy job because the teachers have to know good qualifications of the test.

According to Drs. Wilmar Tinambunan, “Validity, reliability and usability are general considerations in test evaluation that are always important. Next in importance is whether the test measures consistently and accurately.”\(^5\)

---


Well made test can help teacher in increasing teaching-learning process. High quality test items can give the information about how appropriate the students have comprehended the material, which has been given by the teacher. There are some kinds of test. One of them is achievement test.

According to Drs. Wilmar Tinambunan, there are four types of achievement test. Those are placement test, formative test, diagnostic test, and summative test. In this paper, the writer just focuses on the summative test.

The summative test which is made by the English teacher needs to be analyzed whether the test items have been able to measure their students’ achievement or not. The teacher should not draw up the test items carelessly and leaves some of the instructional objectives in English syllabus. In fact, the test can be called as a good test whether the test items are in line with syllabus to measure students’ achievement at the end of learning process.

In June of the 2008/2009 academic year, MTs Annajah administered summative test for the second term. One of the lessons that were tested is English. This test was developed by the English teacher. Nevertheless the content validity of the test items has not been known yet. From this case, the writer is interested in analyzing the content validity of English summative test.

For this analysis, the writer gives title “An Analysis on Content Validity of English Achievement Test Based on the Syllabus for the First Grade of Madrasah Tsanawiyah Annajah Petukangan”.

B. The Limitation of the Problem

The writer limits this paper on the analysis on content validity of English summative test in the first grade of even semester based on the syllabus. Especially focus on indicators in reading skill covering: identifying some information in short functional text in term of instruction, list of things, congratulation, announcement; identifying the topic/main idea in descriptive/procedure text; identifying detail

---

5 Wilmar Tinambunan, Evaluation of..., p. 11.
6 Wilmar Tinambunan, Evaluation of..., p. 7.
information from the text; determining accurate quantifiers in the sentence; identifying rhetorical steps in descriptive/procedure text; identifying communicative purpose in descriptive/procedure text; and denoting linguistic features in descriptive/procedure text.

Indicators on writing skill covering: writing in short functional text in term of instruction, list of things, congratulation, announcement; writing simple sentences; completing descriptive/procedure text; jumbling text/sentence/word into good arrangement; denoting linguistic features in functional text; writing descriptive/procedure text; determining suitable connective of time in the sentence; identifying communicative purpose and rhetorical steps in descriptive/procedure text; determining suitable question tag in the sentence; and using language gambits.

C. The Formulation of the Problem

Based on limitation of the problem, the writer formulates the problem in the question below:

*Does the English summative test for the first grade students of Madrasah Tsanawiyah Annajah Petukangan have high quality on content validity?*

D. The Objective of the Study

The objective of the study is to find empirical evidence if the test items of English summative test carried out at the first grade of Madrasah Tsanawiyah Annajah have high quality on content validity and in line with the school based syllabus.

E. The Method of the Study

Observing and analyzing the English achievement test paper is to know content validity of the test by Comparative Analytic Method that is by comparing content of the test with the English syllabus and the English curriculum for the first grade of junior high school.
F. The Significance of the Study

The results of this study are expected to contribute to the English teachers, the school’s principal and further researchers. The writer also hopes this paper can give meaningful input for teachers of English in developing test items that have content validity. Besides, the result of the study is hoped to enrich the teachers' knowledge of English achievement test and it can be used as a reference for the next researchers who are interested in developing similar study. Especially, for writer, will understand and capable to make a high quality of test items.
CHAPTER II
THEORETICAL FRAMEWORK

This chapter presents and discusses test, validity, achievement test, curriculum, and English curriculum.

A. Test

1. The Meaning of Test

Actually, there are many methods that the teacher can use in evaluating. One of them is test. Before the writer talks more about the test, she would like to elaborate the definition of it. The ‘test’ derives from Latin language “testum” which means a tool to measure the land. In education, test is very important. Test is a device in education and it is a process or device to get the information of students’ achievement in order to achieve the learning objective.

To understand more about test, there are many definitions about the test from the experts. According to Norman E. Gronlund “Test is an instrument or systematic procedure for measuring a sample of behavior (Answer the question “How well does the individual perform-whether in comparison with others or in comparison with a domain of performance tasks”)”. I agree with the statement that test is an instrument to know student’s comprehension.

Michael T. Nietzel said that “Test is systematic procedure for observing and describing a person’s behavior in a standard situation.” It means that test is not only

---


used to score the lesson but also to describe person’s behavior. The student as a test-taker will get the same opportunities in a restricted situation with its instructions.

According to Oxford dictionary, “Test is an examination of a person’s knowledge or ability.”\textsuperscript{9} From the definition, the writer thinks that we can know individual knowledge and ability with using test.

While McNamara states that “Language test is a procedure for gathering evidence of general or specific language abilities from performance on tasks designed to provide a basis for predictions about an individual’s use of those abilities in real world contexts.”\textsuperscript{10} With using test we can know general or specific language abilities of the student and can be used in real world contexts.

Based on the opinion above, the writer can conclude that a test is a procedure or an appliance used to know or measure the students’ comprehension or knowledge about the subject area more accurately. Also test is the way to determine the level of education. A test is a tool or device either in the form of question, order or direction used to measure and value the students’ achievement in mastering the material given by the teacher in accordance to the curriculum.

2. Purposes of Test

There are 8 kinds purpose of test. Those are description, prediction, assessing individual differences, objectives evaluation, domain estimation, mastery decisions, diagnosis, and pre- and post assessment.

It supports by William Wiersma that an illustrative but not exhaustive list of test purposes follows:

a. Description
   Many tests are developed to describe the present status of individuals on a variety of variables.

b. Prediction
   Some tests are used to predict the future performance of the examinees.

c. Assessing individual differences


Some tests are used to differentiate between people in order to identify those who are the highest and those who are the lowest on some measure.

d. Objectives evaluation
   Many schools operate within an objectives based environment. It is essential in such situations to report progress and to plan instruction in terms of the objectives that have been mastered by the student. Of course, the mastery of an objective is a decision that is often based on test scores.

e. Domain estimation
   In some educational setting, there is a well specified domain of content. Many tests are designed to estimate the percentage of a domain that the student understands.

f. Mastery decisions
   Mastery of specific objectives as well as mastery of a larger block of content is often determined on the basis of test scores.

g. Diagnosis
   An instructional plan for a student should be based on a thorough diagnosis of his or her areas of strength and weakness. The educational diagnosis is usually determined by performance on one or more tests. Diagnosis tests are very detailed because they focus on specific and discrete responses of the student.

h. Pre- and post assessment
   These tests need to be sensitive enough to measure the changes in status or performance from one time to another. The focus in this case is the difference between the pretest and the posttest scores rather than the scores themselves.11

Based on the explanation above, it shows that the purposes of test consist of description, prediction, etc.

3. Types of Test and Types of Test Item
   a. Types of Test
      There are many types of test used to measure students’ achievement. Based on its purpose, there are three types of test namely achievement test, proficiency test, and aptitude test.

1) Achievement Tests

---

Achievement test is designed to measure the student’s performance based on the syllabus or program. According to Bill R. Gearheart, “The achievement test attempts to measure the extent to which pupil has achieved in various subject area.”

McNamara states that “Achievement tests are associated with the process of instruction. Examples would be: end of course tests, portfolio assessments, or observational procedures for recording progress on the basis of classroom work and participation. Achievement tests accumulate evidence during, or at the end of, a course of study in order to see whether and where progress has been made in terms of the goals of learning. Achievement tests should support the teaching to which they relate.”

Harrison says, “But the important point which is common to all these situations is that the standard remains constant as far as possible from course to course and from year to year and is external to the individual class or textbook. The conditions for setting and achievement test are that it covers a much wider range of material than a diagnostic test and relates to long-term rather than short-term objectives. This brings up problems of sampling, since what has been learnt in a year (for example) cannot all be assessed in one day, yet the test must reflect the content of the whole course.”

As a conclusion, achievement test is a test to measure the students’ achievement in mastering the past subject area based on the syllabus or program.

2) Proficiency Tests

The second type of test is proficiency tests. This test is used to know the proficiency of test-takers. It is hoped, after giving this test, the test-takers will know their ability in language.

According to McNamara, “Proficiency tests look to the future situation of language use without necessarily any reference to the previous process of teaching.”

---


While Harrison states “The aim of a proficiency test is to assess the student’s ability to apply in actual situations what he has learnt. It seeks to answer the question: ‘having learnt this much, what can the student do with it?’ This type of test is not usually related to any particular course because it is concerned with the student’s current standing in relation to his future needs. In view of this future orientation, a proficiency test is the most suitable vehicle for assessing English for Specific Purposes (ESP), such as language for business, for engineers, or for postgraduate study.”

This test usually consists of the standardized multiple choice items on structure, reading comprehension, listening comprehension, and sometimes on writing.

3) **Aptitude Tests**

The last type of test which the writer would like to discuss is aptitude test. According to J.B. Heaton, “Language aptitude test (or prognostic test) is designed to measure the student’s probable performance in a foreign language which he or she has not started to learn: i.e. it assesses aptitude for learning a language.” Therefore, achievement test measures what the student has learned, while aptitude test measures the student’s ability to learn new study.

Beside that H. Douglas Brown states, “Tests of each of the modes of performance can be focused on a continuum of linguistic units, from smaller to larger: phonology and orthography, words, sentences, and discourse.”

A language aptitude test is designed to measure a person’s capacity or general ability to learn a foreign language and to be successful in that undertaking. Aptitude tests are most often used to measure the suitability of a candidate for a specific program of instruction. Thus, these tests are given before the students begin to study to select them in appropriate section or level of their ability.

b. **Types of Test Item**

---


There are 12 types of test item, such as; questions and answers, true/false, multiple-choice, gap-filling and completion, matching, dictation, cloze, transformation, rewriting, translation, essay, and monologue. The tests are supported by Penny Ur that there are 12 kinds of test item:

1) Questions and Answers
   Simple questions, very often following reading, or as part of an interview; may require short or long answers.
   Example:
   What is the (family) relationship between David Copperfield and Mr. Murdstone?

2) True/False
   A statement is given which is to be marked true or false. This may also be given as a question, in which case the answer is, yes or no.
   Example:
   Addis Ababa is the capital of Egypt.
   Is Addis Ababa the capital of Egypt?

3) Multiple-Choice
   The question consists of a stem and a number of options (usually four), from which the testee has to select the right one.
   Example:
   A person who writes books is called…
   a. a booker  b. an editor
   c. an author  d. a publisher

4) Gap-filling and Completion
   The testee has to complete a sentence by filling a gap or adding something. A gap may or may not be signaled by a blank or dash; the word to be inserted may or may not be given or hinted at.
   Example:
   They (go) to Australia in 1980.
   They_______ to Australia in 1980.

5) Matching
   The testee is faced with two groups of words, phrases or sentences; each item in the first group has to be linked to a different item in the second.
   Example:
   Small     Big
   Many      Little

6) Dictation
   The tester dictates a passage or set of words; the testee writes them down. This mainly tests spelling, perhaps punctuation, and, perhaps surprisingly on
the face of it, listening comprehension: people can only usually write words down accurately from dictation if they understand them. It does not, however, test other writing skills or speech, and involves very little reading.

7) Cloze
Words are omitted from a passage at regular intervals (for example, every seventh word). Usually the first two or three lines are given with no gaps.
Example:
The family is all fine, though Leo had a bad bout of flu last week. He spent most of it lying on the sofa watching________ when he wasn’t sleeping!

8) Transformation
A sentence is given; the testee has to change it according to some given instruction.
Example:
Put into the past tense:
I go to school by bus.

9) Rewriting
A sentence is given; the testee rewrites it, incorporating a given change of expression, but preserving the basic meaning.
Example:
He came to the meeting in spite of his illness. Although…

10) Translation
The testee is asked to translate expressions, sentences or entire passages to or from the target language.

11) Essay
The testee is given a topic, such as ‘Childhood memories’, and asked to write an essay of a specific length.

12) Monologue
The testee is given a topic or question and asked to speak about it for a minute or two.\(^{19}\) This tests oral fluency in ‘long turns’ – something not everyone can do in their mother tongue. It also tests overall knowledge of pronunciation, grammar and vocabulary.

It shows that the test consists of 12 kinds of test item. Those are questions and answers, true/false, etc.

4. The Characteristics of a Good Test

When teacher considers making a test, she or he has to recognize the various tests of each case depending on purpose, time, subject, and the characteristics of a good test. As we know that a test can be said as a good test if it has certain qualifications or certain characteristics, they are: validity, reliability and usability/practicality.

Therefore, the writer would discuss them more detail in the explanations below:

a. **Validity**

  *Validity*. Does the test measure what it is intended to measure? If it does, it is a valid test. Validity is not general but specific. If it tests of vocabulary, does it test vocabulary?

  Test validity will be defined here as the degree to which a test measures what it claims, or purports, to be measuring.\(^{20}\)

  According to Norman, validity refers to the adequacy and appropriateness of the interpretations made from assessments, with regard to a particular use.\(^{21}\)

  In the other hand, Nic said: ‘The general term for how well a test works is ‘validity’, and this too has a variety of meanings.’\(^{22}\)

  Based on William’s opinion, ‘Validity is concerned with the extent to which test measures what it purports to measure and is useful for the purpose for which it was designed.’\(^{23}\)

  According to John, ‘The validity of a test is related to how well the test does what it is supposed to do, namely, to inform us about the examinee’s progress toward


some goal in a curriculum or course of study, or to differentiate levels of ability among various examinees on some task.”

The validity of a test is the extent to which the test measures what it is intended to measure.

From the definitions above, the writer can conclude that there are no differences in the essence of validity, there are only different in the terminology, such as extent and degree and worth while all of them intend to measure the purpose to measure.

Validity refers to the extent to which the results of an evaluation procedure serve particular uses for which they are intended. If the results are to be used to describe pupil achievement, we should like the results to represent the specific achievement we wish to describe, and nothing else. If the results are to be used to predict pupil success in some future activity, we should like them to provide as accurate an estimate of future success as possible.

b. Reliability

The second characteristic of a good test is reliability. A test should be reliable as a measuring instrument. Reliability is the consistency of assessment results. Norman E. Gronlund said that reliability refers to the consistency of assessment results.

John W. Oller states that “The reliability of a test is a matter of how consistently it produces similar results on different occasions under similar circumstances.”

Tests should not be elastic in their measurements: if a student takes a rest at the beginning of a course and again at the end, any improvement in his score should be the result of differences in his skills and not inaccuracies in the test. In the same way, it is important that the student’s score should be the same (or as nearly

---

25 Andrew Harrison, A Language ..., p. 11.
27 John W. Oller, Language Tests..., p. 4.
as possible the same) whether he takes one version of a test or another (for waist measurement, the same result should be obtained whichever tape measure is used) and whether one person marks the test or another (whoever uses the tape measure).

There are therefore three aspects to reliability:
- The circumstances in which the test is taken
- The way in which it is marked
- The uniformity of the assessment it makes.\(^{28}\)

According to J.B. Heaton, Factors affecting the reliability:

1. The extent of the sample of material selected for testing
2. The administration of the test.\(^ {29}\)

The reliability of a test is its consistency. Reliability also means the consistency with which a test measures the same thing all the time.

A test is considered reliable if we get the same result repeatedly. For example, if a test is designed to measure a trait (such as introversion), then each time the test is administered to a subject, the results should be approximately the same. Unfortunately, it is impossible to calculate reliability exactly, but there several different ways to estimate reliability. Reliability does not imply validity. That is, a reliable measure is measuring something consistently, but not necessarily what it is supposed to be measuring.

c. Usability

The third characteristic of good test is practicality or the usability in the preparation of a new test. The term usability, then, refers only to the practicality of the procedure and says nothing about the other qualities present.\(^ {30}\) The teacher must keep in mind a number of very practical considerations which involves economy, ease of administration, scoring and interpretation of result. How long the administering and scoring of test will take, choosing a short test rather longer test.

---

\(^{28}\) Andrew Harrison, *A Language...*, p. 10-11.

\(^{29}\) J. B. Heaton, *Writing English...*, p. 162.

A test’s practicality must be determined in relation to the cost in terms of materials, time, and effort that it requires. This must include the preparation, administration, scoring, and interpretation of the test.\footnote{John W. Oller, Language Tests..., p. 4}

The main questions of practicality are administrative. Just as a teacher cannot be effective without some forward planning (if only to ensure that a cassette recorder will be available and in working order for his lesson), a test must be well organized in advance.\footnote{Andrew Harrison, A Language..., p. 12.}

It should be economical from the viewpoint of both time and money; it should be easily administered and scored; and it should produce results that can be accurately interpreted and applied by available school personnel.\footnote{Robert L. Linn and Norman E. Gronlund, Measurement and..., p. 49.}

Before administering test, some factors about the administration and the test itself must be carefully considered:

1) The availability of enough time for the administration of the test should be fair, because the reliability of test is directly related to the test’s length. If the time allotted to testing is very short, the result of the test tends to be unreliable. A safe procedure is to allot as much time as is necessary to obtain valid and reliable results.

2) The test should be as economical as possible in cost.

3) Any equipment needed during the administration of the test, for example, language laboratory, tape recorder, must be prepared in advance.

4) The length of time needed to get the marking done.

5) The scoring procedure must be appropriate. This is very crucial because it will contribute to the ease of interpretation of the raw scores that can be easily converted into meaningful scores.\footnote{Wilmar Tinambunan, Evaluation of Student Achievement, (Jakarta: Depdikbud, 1988), p. 23.}

In the writer’s opinion the practicality of a test is important in order that test can be administered well. It must be determined in terms of materials, time, and effort that it requires. The writer will not explain all characteristics of good test; therefore it is only about validity because it should be conformed by the title of this paper.
B. Validity

Validity concerns itself with how well the test measures what it aims to measure.\(^{35}\) In this point, the writer will discuss nature of validity and types of validity. When using the term validity in relation to testing and assessment, there are a number of cautions to be borne in mind. Nature of validity is supported by Norman E. Gronlund in the explanations below.

1. Nature of Validity
   a. Validity refers to the appropriateness of the interpretation of the results of an assessment procedure for a given group of individuals, not to the procedure itself.
   b. Validity is a matter of degree; it does not exist on an all-or-none basis. Consequently, we should avoid thinking of assessment results as valid or invalid. Validity is best considered in terms of categories that specify degree, such as high validity, moderate validity, and low validity.
   c. Validity is always specific to some particular use or interpretation. No assessment is valid for all purposes. Assessment results are never just valid; they have a different degree of validity for each particular interpretation to be made.
   d. Validity is a unitary concept. The conceptual nature of validity has typically been described for the testing profession in a set of standards prepared by a joint committee made up of members from three professional organizations that are especially concerned with educational and psychological testing and assessment.
   e. Validity involves an overall evaluative judgment. It requires an evaluation of the degree to which interpretations and uses of assessment results are justified by supporting evidence and in terms of the consequences of those interpretations and uses.\(^{36}\)

It shows that nature of validity consists of 5 cautions to be borne in mind. We sometimes speak of the “validity of a test”, for the sake of convenience, but it is more correct to speak of the validity of the interpretation and use to be made from the results. Thus, when we are appraising or describing validity, it is necessary to

---


consider the specific interpretation or use to be made of the results. Validity is viewed as a unitary concept based on various kinds of evidence.

2. Types of Validity

The most important variable in judging the adequacy of a measurement instrument is its validity. Concerning the types of validity some writers use different labels; however they are similar in procedure and principles.

There are five types of validity. The types of validity are supported by John A. S. Read, the worthwhileness of a test can be projected in these five types of validity. Those are content validity, construct validity, face validity, predictive validity, and concurrent validity.\(^{37}\)

a. Content validity

Content validity talks about the content of test. William Wiersma said, “Content validity is concerned with the extent to which the test is representative of a defined body of content consisting of topics and processes.”\(^{38}\) Therefore, the test should reflect instructional objectives or subject matters. But it is not expected that every knowledge or skills will always appear in the test; there may simply be too many things for all of them to appear in a single test.

Other statement came from John A. S. Read, “Content validity refers to the representativeness of the sample of items or behaviors included in relation to what the test aims to measure. In other words, from a sample of behaviors, the test user wishes to determine how an individual would perform at present in a universe of situations that the test is claimed to represent.”\(^{39}\)

According to Hatch and Farhady, content validity is “the extent to which a test measures a representative sample of the subject matter content.”\(^{40}\)


\(^{40}\) J. Charles Alderson and Brian North, Language Testing in the 1990s: The Communicative
Andrew Harrison defied that content validity is concerned with what goes into the test. The content of a test should be decided by considering the purposes of the assessment, and then draw up as a list known as a content specification. The content specification is important because it ensure as far as possible that the test reflects all the areas to be assessed in suitable proportions and also because it represents a balanced sample, without bias towards the kinds of items which are easiest to write or towards the test material which happens to be available.41

Basically, then, content validity is concerned with how well the sample of test tasks represent the domain of tasks to be measured. Content validity is also called curricular validity, because materials that will be tested are based on curriculum. So, the way to know the content validity is by looking the objective of curriculum and the material of the test.

The explanation above is the requirements of content validity that should be included in making the test in order to get a high quality test.

b. Construct validity

In construct validity, we have to measure the difficulties of the students and also the test has to be qualified. As Bill. R. Gearheart states that construct validity indicates the extent to which a test may be said to measure a particular theoretical construct and trait. Intelligence, anxiety, logical reasoning, and the like represent theoretical constructs.42

In Harrison’s opinion, “Construct validity means how well performance on the assessment can be interpreted as a meaningful measure of some characteristic or quality.”43

John A.S. Read states that, construct validity is evaluated by investigating what qualities a test measures, that is, by determining the degree to which certain explanatory concepts or constructs account for performance on the test. In tests consisting of several subtests, construct validity is supported by the inter correlations of subtests with the total test scores. The emphasis in establishing construct validity should be on the strength of each correlation rather than merely

---

41 Andrew Harrison, A Language…, p. 11.
on its statistical significance. Construct validation aims more at comprehension than at a numerical result.\textsuperscript{44}

c. Face validity

Face validity is a surface or appearance of test. It is supported to Norman E. Gronlund, “Face validity which refers only to the appearance of the assessment.”\textsuperscript{45}

John A.S. Read states that Face validity refers to how the testee views the test in relation to what it aims to measure; it is purely judgmental.\textsuperscript{46}

Then, Harrison also states that “Face validity is concerned with what teachers and students think of the test. Does it appear to them a reasonable way of assessing the students, or does it seem trivial, or too difficult, or unrealistic? The only way to find out about face validity is to ask the teachers and students concerned for their opinions, either formally by means of a questionnaire or informally by discussion in class or staff room.”\textsuperscript{47}

If a test looks and acts like to measure the behavior, attitudes, skills or knowledge it is supposed to assess, then we say the test has face validity or a test is said to have face validity if it looks as if measures what it is supposed to measure. For example a test pretended to measure pronunciation ability but which did not require the candidate to speak (and there have been some) it can be thought that is lack of face validity.

In the other hand, the writer is in the opinion that face validity is the construction or the feature of the tests, they are: spelling is how the word is printed correctly; marker/punctuation mark is the symbol used in punctuating sentences; letters is a mark expressing speech sound; printing, how the printing of the test, clear or no; and composition is the arrangement of the test.

\textsuperscript{44} A. S. Read, \textit{Papers on…}, p. 2.


\textsuperscript{47} Andrew Harrison, \textit{A Language…}, p. 11.
d. Predictive validity

According to John A.S. Read, “Predictive validity is demonstrated by how well the test scores correlate with a criterion measure taken at a much later date. It is computed when the test user wishes to forecast an individual’s future standing on some variable of particular significance that is different from the test.”

In addition, Bill R. Gearheart states that predictive validity involves testing the effectiveness of a test against future performance in the areas purportedly measured by the test. Aptitude test are often validated by checking on the future performance of individuals who supposedly have “high” or “low” predicted success is a given specialty.

For example:

<table>
<thead>
<tr>
<th>September 17</th>
<th>December 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scholastic Aptitude scores (test performance)</td>
<td>Achievement Test scores (criterion performance)</td>
</tr>
</tbody>
</table>

e. Concurrent validity

According to Bill R. Gearheart, “Concurrent validity is determined by comparing test performance and some criterion data which is available at the time of testing. Concurrent validity should be used with tests employed to diagnose existing status, but is not as appropriate for use in predicting future outcomes.”

Concurrent validity is demonstrated by how well the test scores compare with one or more external variables considered to provide a direct measure of the characteristic or behavior in question. A concurrent validity coefficient is computed when the test user wishes to estimate an individual’s present standing on some variable of particular significance that is different from the test.

For example:


51 John A. S Read, Papers on…, p. 2.
achievement Test

The purpose of achievement test, as its name reflects, is to establish how successful individual students, groups of students, or the courses themselves have been in achieving objectives of language courses.\(^{52}\)

An achievement tests are usually a formal examination given at the end of the school year or at the end of the course. The achievement tests may be written and administered by ministries of education, official examining boards, or by members of teaching institution.

1. The Construction of an Achievement Test

There are four things in the construction of an achievement test. As Louis J. Karmel said that the construction of an achievement test entailed a careful analysis of the field to be examined.

a. The reasons for the construction of the instrument must be clearly evident.

b. An exhaustive and definitive outline of the subject matter to be used is made.

c. The reasons for construction and an outline of the content are reviewed with such specialists as classroom teachers, educators, and test makers.

d. To compose test items for each part of the content outline, and then ask representative educators to comment on their importance, clarity of expression, and representativeness of subject matter.\(^{53}\)

It shows that the construction of an achievement test is included in four fields to be examined.

2. Basic Principles of Achievement Testing

In order to have a good achievement test form, a test maker should consider that achievement test must be constructed well by paying attention

---


to some following basic principles. Wilmar Tinambunan said in his book that the following principles of achievement testing provide a firm base for constructing and using classroom test as a positive force in the teaching learning process.

a. Achievement test should measure clearly defined learning outcomes that are in harmony with the instructional objectives.
b. Achievement test should measure an adequate sample of the learning outcomes and subject-matter content included in the instruction.
c. Achievement tests should include the types of test items which are most appropriate for measuring the desired learning outcomes.
d. Achievement tests should be designed to fit the particular uses to be made of the results.
e. Achievement test should be made as reliable as possible and should be then interpreted with caution.
f. Achievement test should be used to improve student learning.\(^\text{54}\)

It shows that the basic principles of achievement testing are included in six analyses to be examined.

3. Types of Achievement Test

According to Wilmar Tinambunan, there are many types of test used to measure students’ achievement. Based on its function, there are four types of achievement test which are commonly used by teachers in the classroom, namely placement test, formative test, diagnostic test, and summative test.\(^\text{55}\)

a. Placement Tests

To determine pupil performance at the beginning of instruction uses placement tests. As Norman E. Gronlund said that placement evaluation is to determine pupil performance at the beginning of instruction.\(^\text{56}\)

Placement test is designed to sort new students into teaching groups, so that they can start a course at approximately the same level as the other students in the class. It is concerned with the student’s present standing, and so relates to general ability rather than specific points of learning. As a rule the results are needed quickly so that teaching may begin. This puts a severe constraint upon the types of test that can be used. At the same time a variety of tests is

\(^{54}\) Wilmar Tinambunan, *Evaluation of...,* p. 28-35.

\(^{55}\) Wilmar Tinambunan, *Evaluation of...,* p. 7-8

necessary because a range of different activities is more likely to give an accurate overall picture of a student’s level than a single assessment. The goal of placement evaluation is to determine the position in the instructional sequence and the mode of instruction that is most beneficial for each pupil.

b. Formative Tests

Formative test is used at the end of a unit in the course book or after a lesson designed. Formative evaluation is to monitor learning progress during instruction. Its purpose is to provide continuous feedback to both pupil and teacher concerning learning successes and failures.

Formative evaluation depends heavily on specially prepared tests for each segment of instruction (e.g., unit, chapter).

c. Diagnostic Tests

Diagnostic evaluation is to diagnose learning difficulties during instruction. Diagnostic test checks on student’s progress in learning particular elements of the course. It is used for example at the end of a unit in the course book or after a lesson designed to teach one particular point. The diagnostic test tries to answer the question: ‘How well have the students learnt this particular material?’ Since it relates to particular elements in the course which have just been taught, the assessment will give immediate feedback to the student.

Diagnostic evaluation is concerned with the persistent or recurring learning difficulties that are left unresolved by the standard corrective prescriptions of

---

57 Andrew Harrison, A Language ..., p. 4.
62 Andrew Harrison, A Language ..., p. 6.
formative evaluation. The main aim of diagnostic evaluation is to determine the causes of persistent learning problems and to formulate a plan for remedial action.\(^\text{63}\)

d. Summative Tests

Summative evaluation is to evaluate achievement at the end of instruction.\(^\text{64}\) Summative evaluation typically comes at the end of a course (or unit) of instruction. It is designed to determine the extent to which the instructional objectives have been achieved and is used primarily for assigning course grades or for certifying pupil mastery of the intended learning outcomes. Although the main purpose of summative evaluation is grading, or the certification of pupil achievement, it also provides information for judging the appropriateness of the course objectives and the effectiveness of the instruction.\(^\text{65}\)

It shows that the kinds of achievement test are divided into 4 types. Those are placement test, formative test, diagnostic test, and summative test.

The construction of a standardized achievement test is to state with clarity the reasons and objectives behind the instrument. It involves an exhaustive and definitive outline of the subject matter to be tested. In order to have a good achievement test form, a test maker should consider that achievement test must be constructed well by paying attention to some following basic principles.

There are many types of test used to measure students’ achievement. While there are a number of tests that teachers usually carry out in the classroom, however they can be used based on their functions. The writer presents only one of them which directly related to the analysis written in this paper. It is summative test.

D. Curriculum

1. Meaning of Curriculum


The meaning of curriculum in narrower sense: “Curriculum is one of those terms, like communicative, authentic and many others in teaching, that are used in a confusing variety of senses.”

The meaning of curriculum in broad sense: “The curriculum includes the goals, objectives, contents, processes, resources and mean of evaluation of all the learning experiences planned for pupils both in and out of school.”

According to Oxford dictionary, curriculum is subjects included in a course of study or taught in a school, college, etc.

If we look at the individual elements within these definitions, we see that discussions about the curriculum can include discussions about aims/objectives, content, methods/learning activities, books/materials, and assessment/evaluation.

2. Purpose of Education

An analysis of the purposes of education which has influenced many other writers is Skilbeck’s categorization of the ‘value systems’ underlying three educational traditions.

Skilbeck’s three traditions are:

a. Classical humanism

   In this tradition the main purpose of education is to transmit valued knowledge and culture to an elite section of the next generation and, in so doing, to develop their general intellectual abilities. The curriculum is determined mainly by the valued subject content, which exists outside the learners and should be transmitted to them.

b. Reconstructionism

   In this tradition the main purpose is to bring about desired social change. In order to achieve this, the focus shifts onto providing every individual with knowledge and skills that are useful for social life. The curriculum is carefully planned around taxonomies of objectives which each learner should be enabling to master.

c. Progressivism

   The main purpose is to enable each individual to develop towards self-fulfillment. Since self-fulfillment takes different form for different people, the


focus is on nurturing natural growth processes rather than planning endpoints.\footnote{Roger Bowers and Christopher Brumfit, \textit{Applied Linguistics}..., p. 14.}

Classical humanism is exemplified by the traditional ‘grammar school’ system of Britain and many other European countries; reconstructionism, by the detailed specification of learning objectives that now dominates many school syllabuses; and progressivism, by the emphasis on freedom and creativity in the school founded by Waldorf and Pestalozzi.

3. Approach of Curriculum

In recent years in foreign language teaching, we have seen changing emphases in our conceptions of the goals of learning, existing either at different times or at the same time but in different places. In this section, we will set out, by way of comparison, the dominant features of three approaches to the curriculum.

a. The grammar-based curriculum

1) The main goal of language teaching is to enable learners to master the grammar and vocabulary of the language. (A secondary goal may be to pass on knowledge of the country and its culture).

2) The objectives are mainly defined with reference to individual structures or items of vocabulary.

3) The syllabus attempts to select and sequence these structures and vocabulary, using criteria such as complexity, importance, teach ability and so on.

4) The materials provide learners with examples of language structures and vocabulary in texts devised specially for the purpose.

5) The classroom activities provide learners with opportunities to understand and use the language forms as accurately as possible.

b. The function-based curriculum

1) The main aim is to equip learners to fulfill their communicative needs is an appropriate range of situations.

2) The objectives are defined mainly in behavioral terms: expressing or understanding particular communicative functions or notions, acquiring useful skills, and so on.

3) The syllabus selects and sequences these functions or skills according to criteria such as usefulness, complexity of the language they require, and so on.

4) The materials provide examples of language being used for a variety
of communicative purposes. (The materials are often authentic texts taken from real situations).
5) The classroom activities provide learners with opportunities to practice conveying and understanding meanings.

c. The process-based curriculum
1) The main aim is to create contexts which will stimulate the potential for natural language growth.
2) The objectives are not defined in terms of detailed behavior which every learner is expected to achieve but in non-language terms e.g. topics, tasks or problems to be solved.
3) The syllabus provides a sequence of contexts for learning which are roughly graded according to the demands they make on communication skills rather than by strict linguistic criteria.
4) The materials provide a focus for using language in order to exchange meanings about these topics or tasks.
5) The learning activities consist mainly of the use of language for the purposes of communication. There is a minimum of language correction and of form-oriented practice.

As an example of one approach to balancing the different aims and aspects of language learning, I will take the ‘multi-level’ or ‘variable focus’ curriculum which is described by Allen and presented in an adapted form in Stern:69

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural</td>
<td>Functional</td>
<td>Experiential</td>
</tr>
<tr>
<td>Focus on language (formal features)</td>
<td>Focus on language (discourse features)</td>
<td>Focus on the use of language</td>
</tr>
<tr>
<td>a. Structural control</td>
<td>a. Discourse control</td>
<td>a. Situational or topical control</td>
</tr>
<tr>
<td>b. Materials simplified structurally</td>
<td>b. Materials simplified functionally</td>
<td>b. Authentic language</td>
</tr>
</tbody>
</table>

Allen’s proposal is that the curriculum should contain all three levels all the time but that the emphasis can change at different stages of learning. For example, in the early stages the emphasis might be on acquiring a structural foundation (level 1); the focus may then shift onto the ways in which communication exploits these

---

structures (level 2); then the focus may shift onto creating contexts where the language can be used for real interaction (level 3).

E. English Curriculum

After we talked about curriculum in general, now, the writer will discuss English curriculum. There are some points will be discussed like: Background, Competence Model, Purpose, Scope, Competency Standard and Basic Competency.

1. Background

Language has a central role in student development of intellectual, social, and emotional and it is one of supporting success in learning all subjects. Learning language is hoped helping student to know him or her self, him or her culture, and other culture. Beside that, learning language also helps student being ability in expressing idea and feeling, participating in the society, and even finding along with using analytical and imaginative ability in him or her self.

English is a tool for communication orally and written. Communication is understand and expressing the information, thinking, feeling, and developing knowledge, technology, and culture. Communication ability in real meaning is discourse competence, it is ability to comprehend and/or produce orally and/or written text, which is realized in four language skills, they are listening, speaking, reading and writing. These four skills are used to perceive or create discourse in live society. Therefore, English subject is aimed to develop those skills, so that the graduate can communicate and be discourse with English in certain literacy stage.

Literacy stage include performative, functional, informational, and epistemic. In performative stage, the person can read, write, listen, and speak with using symbols. In functional stage, the person can use language to fulfill daily needs like reading newspaper, manual or instruction. In informational stage, the person can get knowledge with language ability, whereas in epistemic stage, the person can express knowledge to target language (Wells, 1987).

English learning in SMP/MTs is aimed to reach functional stage, like communicating orally and written to solve daily problems, but in SMA/MA is hoped
reaching informational stage because they are prepared to continue their education to university. In epistemic stage is assumed that it is to be reach by student in SMA/MA because English in Indonesia is as a foreign language.\(^70\)

2. Competence Model

There are so many competence models which relate to language area that see linguistic competence from various perspectives so far. In this curriculum, the linguistic competence model, which is used, is model that is motivated by language pedagogy considerations which have developed or have been revolution since Canale and Swaim model for about 30 years ago.

One of the newest model, which is in language education literature is faced by Celce-Murcia, Dornyei and Thurrell (1995) that is suitable with theoretical view that language is communication, not only sets of rule. The implication of linguistic competence model which is formulated is the model that prepare student to use language to participate in language society. This model is formulated as Communicative Competence or “Kompetensi Komunikatif (KK)” which is represented by Celce-Murcia et al. (1995:10), like:

Picture 2.1: Communicative Competence Model (by Celce-Murcia et al)\textsuperscript{71}

Schematic representation in the picture 2.1 shows that the major competence which is purposed with language education is Discourse Competence or “Kompetensi Wacana (KW)”. If someone communicate orally or written, it means he/she is on a discourse.

According to McCarthy and Carter (2001:88), Participating in conversation, reading and writing automatically will activate discourse competence which means using any strategy or procedure to bring about something in language element, structure, and semantics.

Discourse competence will be required if students get supporting competence, like: linguistic competence, actional competence, sociocultural competence and strategic competence.\textsuperscript{72}

3. Purpose

Based on the curriculum 2006, English subject in Junior High School is intended to the students to have abilities like:

4. To develop communicating competence orally and written to reach functional literacy.
5. To have an awareness of English essence and important to enhance the national potency in global society.
6. To develop understanding of students about the relevancy between language and culture.\textsuperscript{73}

4. Scope

The scopes of English subject in SMP/MTs include:


\textsuperscript{73} PUSKURNAS, \textit{Standar Isi dan Standar Kompetensi…}, p. 278.
a. Discourse competence is ability to comprehend and/or produce orally and/or written text that is realized in four language skills, they are listening, speaking, reading, and writing integrated to reach functional literacy stage;

b. Ability to comprehend and create any kinds of short functional text and monologue along with essay has the form of procedure, descriptive, recount, narrative, and report. The material graduation is visible in using vocabulary, grammar, and rhetoric steps;

c. Supporting competence is linguistic competence (using grammar and vocabulary, phonetics, syntax), socio cultural competence (using expression, and language action receive in any communication context), strategic competence (solving the problem which is appear in communication process with any ways, in order that communication goes on), and actional competence (using instrument of discourse form).

5. Competence Standard and Basic Competency

The following table describes the competence standard and basic competency of reading and writing skills in the school level of curriculum. We can see the table follow:

<table>
<thead>
<tr>
<th>Standar Kompetensi</th>
<th>Kompetensi Dasar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Membaca</td>
<td></td>
</tr>
<tr>
<td>11. Memahami makna teks tulis fungsional dan esei pendek sangat sederhana berbentuk descriptive dan</td>
<td>11.1 Merespon makna yang terdapat dalam teks tulis fungsional pendek sangat sederhana secara akurat, lancar dan berterima yang berkaitan dengan lingkungan terdekat</td>
</tr>
<tr>
<td>11.2 Merespon makna dan langkah retorika secara</td>
<td></td>
</tr>
</tbody>
</table>

---

74 PUSKURNAS, *Standar Isi dan Standar Kompetensi...,* p. 278.
procedure yang berkaitan dengan lingkungan terdekat

akurat, lancar dan berterima dalam esei sangat sederhana yang berkaitan dengan lingkungan terdekat dalam teks berbentuk descriptive dan procedure

11.3 Membaca nyaring bermakna teks fungsional dan esei pendek dan sangat sederhana berbentuk descriptive dan procedure dengan ucapan, tekanan dan intonasi yang berterima

Menulis

12. Mengungkapkan makna dalam teks tulis fungsional dan esei pendek sangat sederhana berbentuk descriptive dan procedure untuk berinteraksi dengan lingkungan terdekat

12.1 Mengungkapkan makna dalam teks tulis fungsional pendek sangat sederhana dengan menggunakan ragam bahasa tulis secara akurat, lancar, dan berterima untuk berinteraksi dengan lingkungan terdekat

12.2 Mengungkapkan makna dan langkah retorika dalam esei pendek sangat sederhana dengan menggunakan ragam bahasa tulis secara akurat, lancar dan berterima untuk berinteraksi dengan lingkungan terdekat dalam teks berbentuk descriptive dan procedure

Based on Competence Standard and Basic Competency above, we will analyze content validity of the test items that is made by Annajah Junior High School. Beside from Competence Standard and Basic Competency, the analysis is also determined by indicators in reading and writing skills.
CHAPTER III
THE IMPLEMENTATION OF THE RESEARCH

This chapter presents and discusses the methodology of the research, the findings of the research and the discussions.

A. The Methodology of the Research

1. The Objective of the Research

From this case, the writer is interested to focus is to find empirical evidence if the test items of English summative test carried out at the first grade of Madrasah Tsanawiyah Annajah have included content validity and in line with the school based syllabus.

Then, the writer wants to know the information from the English teacher as a test designer, how to construct the good English test and to look for the reasons that there is still some inappropriateness in constructing the English test.

2. The Time and Location

The research was done by the writer from September 4th, 2009 to January 28th, 2010. It was conducted at the first grade students of Madrasah Tsanawiyah Annajah, which is located on Jalan Ciledug Raya Petukangan Selatan Pesanggrahan, South Jakarta.

3. The Techniques of Data Collecting

The next step of this research is collecting the data. The sources of data are reading and writing test of English summative test. The instruments of the data are question sheets, blue print of the questions, English syllabus, English curriculum, and the English teacher as a test designer for the first grade students of Madrasah Tsanawiyah Annajah to be further analyze. Beside that, the writer also did the
interview with the English teacher. English summative test was held by school on June 19th, 2009 for the first grade students of Madrasah Tsanawiyah Annajah Petukangan.

4. The Techniques of Data Analysis

The collected data are quantitatively and qualitatively analyzed. Quantitatively, the data are calculated using simple percentage formula. It is the formula:

\[ P = \frac{f}{n} \times 100 \%
\]

**Note:**
- \( P \) = Percentage
- \( f \) = Frequency
- \( n \) = Number / Amount

It is used to see how many percent the test covers the instructions of the curriculum. Qualitatively, the test items are studied in terms of their appropriateness to the English syllabus and the themes.

In addition, the writer also compares the percentage with the criteria adopted from Arikunto’s opinion:

<table>
<thead>
<tr>
<th>Criteria of Test Result Percentage</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>76 – 100 %</td>
<td>Good</td>
</tr>
<tr>
<td>56 – 75 %</td>
<td>Sufficient</td>
</tr>
<tr>
<td>40 – 55 %</td>
<td>Fair</td>
</tr>
</tbody>
</table>

---

To increase our insight, the writer did the interview with the English teacher. The interview was done to know more information how to make a good English test.

This interview was conducted on Thursday, January 28th, 2010 at 10.00 in the English teacher’s office and the research informant was the English teacher in MTs Annajah. It also was conducted to confirm the data collected from English summative test. The interview can be seen in the appendices clearly.

B. The Findings of the Research and the Discussions

1. The Description of the Data

The writer analyzed the test items whether they are in line with the latest English syllabus administrated at Madrasah Tsanawiyah Annajah. The data analyzes such as question sheets, blue print of the questions, English syllabus, English curriculum, and the English teacher as a test designer for the first grade students of Madrasah Tsanawiyah Annajah to be further analyze (see appendices).

The data that the writer used in this study is the English summative test for even semester, which is called as “Evaluasi Hasil Belajar”, the 2008/2009 academic year for the first grade students of Madrasah Tsanawiyah Annajah Petukangan. Based on regulation about the standard of educational evaluation established by minister of National Education no. 20, 2007, the summative test of particular semester must be constructed based on the some indicators which should be achieved at that semester. Because of that, the summative test of the first grade students of Madrasah Tsanawiyah Annajah Petukangan was constructed based on particular indicators.

The total number of the test items is 45 items which consist of 40 multiple choice and 5 essay test items. The test was held on Friday, June 19th 2009 with the given time 90 minutes. The test was developed by the English teacher in Madrasah Tsanawiyah Annajah Petukangan.
The following table describes the total frequency of the appropriateness and the inappropriateness of English summative test items with indicators in the latest English syllabus. Based on the data of items analysis result, we can see the table as follow (see appendices, page. 64-67):

Table 3.2
The Appropriateness of English Summative Test with the Indicators in the Latest English syllabus at Even Semester of the First Grade Students at MTs Annajah

<table>
<thead>
<tr>
<th>No</th>
<th>Based on Indicators in the English Syllabus</th>
<th>Items Number</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td><strong>Reading</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Identifying some information in short functional text in term of instruction, list of things, congratulation, announcement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Identifying the topic/main idea in descriptive/procedure text</td>
<td>1, 29</td>
<td>2 Items</td>
</tr>
<tr>
<td>3.</td>
<td>Identifying detail information from the text</td>
<td>2, 3, 14, 15, 27, 41</td>
<td>6 Items</td>
</tr>
<tr>
<td>4.</td>
<td>Determining suitable quantifiers in the sentence</td>
<td>21, 22</td>
<td>2 Items</td>
</tr>
<tr>
<td>5.</td>
<td>Identifying rhetorical steps in descriptive/procedure text</td>
<td>17, 30</td>
<td>2 Items</td>
</tr>
<tr>
<td>6.</td>
<td>Identifying communicative purpose in descriptive/procedure text</td>
<td>28</td>
<td>1 Item</td>
</tr>
<tr>
<td>7.</td>
<td>Denoting linguistic features in descriptive/procedure text</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td><strong>Writing</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Writing in short functional text in term of instruction, list of things, congratulation, announcement</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>Writing simple sentences</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>Completing descriptive/procedure text</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>Jumbling text/sentence/word into good arrangement</td>
<td>7, 35</td>
<td>2 Items</td>
</tr>
<tr>
<td>12.</td>
<td>Denoting linguistic features in functional text</td>
<td>8, 9, 44</td>
<td>3 Items</td>
</tr>
<tr>
<td>13.</td>
<td>Writing descriptive/procedure text</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14.</td>
<td>Determining suitable connective of time in the sentence</td>
<td>34</td>
<td>1 Item</td>
</tr>
<tr>
<td>15.</td>
<td>Identifying communicative purpose and rhetorical steps in descriptive/procedure text</td>
<td>45</td>
<td>1 Item</td>
</tr>
<tr>
<td>16.</td>
<td>Determining suitable question tag in the sentence</td>
<td>12</td>
<td>1 Item</td>
</tr>
<tr>
<td>17.</td>
<td>Using language gambits</td>
<td>4, 5, 6, 10, 11, 18, 19, 20, 23, 24, 25, 31, 32, 33, 39, 40, 43</td>
<td>17 Items</td>
</tr>
</tbody>
</table>

**Total** 38 Items

Based on the table above, there are many indicators in the competence standard of reading and writing, but the test just covers few of them. In reading, those are identifying the topic/main idea in descriptive/procedure text; identifying detail information from the text; determining suitable quantifiers in the sentence; identifying rhetorical steps in descriptive/procedure text; identifying communicative purpose in descriptive/procedure text.

In writing, there are jumbling text/sentence/word into good arrangement; denoting linguistic features in functional text; determining suitable connective of time in the sentence; identifying communicative purpose and rhetorical steps in
descriptive/procedure text; determining suitable question tag in the sentence, and using language gambits.

In the other word, the indicators: identifying some information in short functional text in term of instruction, list of things, congratulation, announcement; and denoting linguistic features in descriptive/procedure text, in reading and the indicators: writing in short functional text in term of instruction, list of things, congratulation, announcement; writing simple sentences; completing descriptive/procedure text; and writing descriptive/procedure text, in writing are not covered in the test.

Beside that, there are also some of expansion items number which are not proportional to the test.

Table 3.3
The Inappropriateness of English Summative Test with the Indicators in the Latest English syllabus at Even Semester of the First Grade Students at MTs Annajah

<table>
<thead>
<tr>
<th>No</th>
<th>Based on Indicators in the English Syllabus</th>
<th>Items Number</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Identifying the kinds of text that are meant by the test</td>
<td>16, 42,</td>
<td>2 Items</td>
</tr>
<tr>
<td>2</td>
<td>Determining suitable word meaning</td>
<td>36,</td>
<td>1 Item</td>
</tr>
<tr>
<td>3</td>
<td>Determining suitable vocabulary</td>
<td>13, 26, 37, 38,</td>
<td>4 Items</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td>7 Items</td>
</tr>
</tbody>
</table>

From the table above, we know that there are three indicators which are not appropriate to the English syllabus. They are identifying the kinds of text that are meant by the test, determining suitable word meaning, and determining suitable vocabulary. They cover 7 items number.

2. The Analysis of the Data

In this section, the writer analyzed the collected data qualitatively based on Competence Standard. On the previous table, there are 38 items which are
appropriate with the indicators in the English syllabus. On the other hand, there are also 7 items that are not appropriate with the indicators in the English syllabus. In the table below, the writer will explain the items, one by one, based on indicators reading and writing.

**Table 3.4**

**Based on Competence Standard: Reading**

"Memahami makna teks tulis fungsional dan esai pendek sangat sederhana berbentuk descriptive dan procedure yang berkaitan dengan lingkungan terdekat”

<table>
<thead>
<tr>
<th>No</th>
<th>Item</th>
<th>Questions and Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td><strong>Temple</strong>&lt;br&gt;There are many temples in Indonesia. The biggest Budha temple is Borobudur. It lies in Magelang Central Java. It is one of the wonderful buildings in the world.&lt;br&gt;There is a beautiful temple in Prambanan. It is called Prambanan temple. It is Hindu temple lies in Plaosan village. Many tourists from foreign countries go there.&lt;br&gt;The other temples are Bima temple, Arjuna temple, Srikandi temple, Puntadewa temple, Sembadra temple, and Semar temple. All the temples lie on the high land of Dieng. All of them are not as big as Prambanan temple.&lt;br&gt;Paragraph one is talking about…&lt;br&gt;a. the temple in the high land of Dieng c. Prambanan temple&lt;br&gt;a. the biggest Budha temple d. Hindu temple</td>
<td>Appropriate with indicator in English syllabus:&lt;br&gt;This question is intended to measure students’ ability to know the topic in the paragraph one. Thus, this item is appropriate to the recommended indicator, namely, Identifying the topic/main idea in descriptive/procedure text.</td>
</tr>
</tbody>
</table>
| 2. | **Question:**

“It” is Hindu temple… (Paragraph two) refers to…

a. Borobudur temple c. Plaosan temple |
b. Prambanan temple  
d. Land of Dieng

**Appropriate with indicator in English syllabus:**
The question asks about the meaning of pronoun in paragraph two. So, the question is intended to measure students’ ability which is suitable with the indicator: Identifying detail information from the text.

**Question:**
Semar temple lies in the...

- a. World
- b. Magelang Central
- c. Plaosan village
- d. high land of Dieng

**Appropriate with indicator in English syllabus:**
The question wants to know students’ ability in determining specific information. It is appropriate with the indicator, namely, Identifying detail information from the text.

**Question:**
I have a house. My house is in a village. It is not big but beautiful. There is a small garden behind it. My house has some rooms. They are a living room, three bedrooms, a kitchen, a bathroom and a dining room.

I have a table and five chairs in my living room. I also have a television set in it. My bedroom is small but nice. I have a small desk, a chair, and a shelf in it. There are a lot of books in the shelf. I like my house very much.

What is behind the house?

- a. a small garden
- b. a village
- c. a bedroom
- d. a living room

**Appropriate with indicator in English syllabus:**
The item is intended to measure students’ ability in looking for specific information from the text. So that, it is suitable with the indicator: Identifying
<table>
<thead>
<tr>
<th>Question: Where is the television set? In the…</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. dining room</td>
</tr>
<tr>
<td>b. living room</td>
</tr>
</tbody>
</table>

**Appropriate with indicator in English syllabus:**
The item asks about located of the television set. This is appropriate with the indicator, Identifying detail information from the text.

<table>
<thead>
<tr>
<th>Question: What kind of text is the text above?</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Descriptive</td>
</tr>
<tr>
<td>b. Procedure</td>
</tr>
</tbody>
</table>

**Inappropriate with indicator in English syllabus:**
This question is intended to know the kinds of text. So, Identifying the kinds of text that are meant by the test is not suitable with the indicator.

<table>
<thead>
<tr>
<th>Question: The first paragraph is called…</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Orientation</td>
</tr>
<tr>
<td>b. Identification</td>
</tr>
</tbody>
</table>

**Appropriate with indicator in English syllabus:**
This question is intended to measure students’ ability in determining the generic structure of the descriptive text. It is appropriate to the indicator, Identifying rhetorical steps in descriptive/procedure text.

<table>
<thead>
<tr>
<th>Question: Mrs. Frida needs… of eggs to make a cake.</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. one fourth kilograms</td>
</tr>
<tr>
<td>b. a bottle</td>
</tr>
</tbody>
</table>

**Appropriate with indicator in English syllabus:**
This question asks how many eggs that Mrs. Frida need to make a cake. It is suitable with the suggested indicator, namely, Determining suitable
quantifiers in the sentence.

<table>
<thead>
<tr>
<th>Question:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>There are… books in the library, but there only a few books in the office.</td>
<td></td>
</tr>
<tr>
<td>a. a lot of</td>
<td>c. much</td>
</tr>
<tr>
<td>b. a few</td>
<td>d. a little</td>
</tr>
</tbody>
</table>

**Appropriate with indicator in English syllabus:**
The item asks about the amount of books in the library and in the office. It is very appropriate with the indicator, Determining suitable quantifiers in the sentence.

<table>
<thead>
<tr>
<th>Question:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DOUBLE CHOCOLATE CAKE</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Ingredients:</strong></td>
<td></td>
</tr>
<tr>
<td>30 ml shortening, 60gr unsweetened chocolate, 250 ml sugar, an egg, 375 ml sifted all – purpose flour, 5 ml baking soda, 1 ml salt, 250 ml buttermilk or soured milk, 5 ml vanilla, creamy Chocolate Icing.</td>
<td></td>
</tr>
<tr>
<td><strong>Direction:</strong></td>
<td></td>
</tr>
<tr>
<td>1. Heat oven to 175 Celsius degrees.</td>
<td></td>
</tr>
<tr>
<td>2. Grease a 9 – inch square cake pan.</td>
<td></td>
</tr>
<tr>
<td>3. Combine shortening and chocolate in a small saucepan and set over lowest heat to melt.</td>
<td></td>
</tr>
<tr>
<td>4. Combine sugar, egg, and chocolate mixture in mixing bowl and beat well.</td>
<td></td>
</tr>
<tr>
<td>5. Sift flour, baking soda and salt together, and add to sugar mixture of buttermilk or soured milk and vanilla.</td>
<td></td>
</tr>
<tr>
<td>6. Pour into prepared pan and bake about 35 minutes or until top springs back when touched lightly in centre.</td>
<td></td>
</tr>
<tr>
<td>7. Cool in pan, ice with creamy chocolate icing.</td>
<td></td>
</tr>
</tbody>
</table>

Which sentences is **INCORRECT** in baking double chocolate cake?
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. We need 375 ml sifted all purpose flour</td>
<td></td>
</tr>
<tr>
<td>b. The fourth step is to combine sugar, egg, and chocolate</td>
<td></td>
</tr>
</tbody>
</table>
c. Cool in pan, ice with creamy chocolate icing  
d. Combine shortening and chocolate, set over highest heat to melt

**Appropriate with indicator in English syllabus:**  
The item look for the incorrect information from the text. So, the question is appropriate with the indicator: Identifying detail information from the text.

**Question:**  
What is the purpose of the text? It is…  
- a. to inform to do or make something  
- b. to describe a particular thing  
- c. to amuse the readers  
- d. to tell the event

**Appropriate with indicator in English syllabus:**  
This question is intended to measure students’ ability in determining the purpose of text. It is appropriate with the indicator, namely, Identifying communicative purpose in descriptive/procedure text.

**Question:**  
What does the text tell us about?  
- a. How to beat the dough  
- b. The process of baking a chocolate cake  
- c. The process of mixing the ingredients  
- d. How to sift the flour

**Appropriate with indicator in English syllabus:**  
This question is intended to measure students’ ability to know the topic in the text. Thus, this item is appropriate to the recommended indicator, namely, Identifying the topic/main idea in descriptive/procedure text.

**Question:**  
What is the general structure of the text? It is…  
- a. aim – orientation – event  
- b. orientation – complication  
- c. aim – material – steps  
- d. general classification – description

**Appropriate with indicator in English syllabus:**
This question is intended to measure students’ ability in determining the generic structure of the procedure text. It is appropriate to the indicator, Identifying rhetorical steps in descriptive/procedure text.

### Question:
It is **fixed price** in Supermarket. What does the underlined word mean?

- a. harga pas
- b. harga diskon
- c. menawar
- d. harga tawaran

#### Inappropriate with indicator in English syllabus:
The question asks about the meaning in the underline words. This item is not appropriate to the indicator in English syllabus. Thus, Determining suitable word meaning is not suitable.

### Question:
A : What are you going to do?
B : I am going to buy a pen in the…
A : Would you buy an eraser for me, please?
B : Certainly.

- a. stationers
- b. shoe store
- c. green grocer
- d. grocery store

#### Inappropriate with indicator in English syllabus:
The dialogue asks about the right kinds of shop. This item is not suitable because the indicator, Determining suitable vocabulary is not appropriate to the English Syllabus.

### Question:
Which ingredient is needed in making fried rice, except?

- a. Noodles
- b. Rice
- c. Salt
- d. Cucumber

#### Inappropriate with indicator in English syllabus:
The item asks about things that are needed in making fried rice. This item is not suitable because the indicator, Determining suitable vocabulary is not
Our Living Room

Our living room is not big but comfortable. It has one table and four chairs. There is a vase of flowers on the table. A clock is hanging on the wall. To the right of the clock, there is a calendar. On the floor, in front of the television set, there is a rug with two cushions on it. In the corner of the room, there is a small table with a cassette player and some cassettes on it.

What is the writer’s living room like?

Questions and Explanation

A: Do you get up at 4 every morning?
B: ...

A: What time do you get up?
B: I get up at 4.30 every morning.

a. No, I don’t get up at 4.
   c. No, she doesn’t get up at 4.
b. Yes, I get up at 4
   d. Yes, she gets up at 4.

Appropriate with indicator in English syllabus:
The question is about dialogue that tells about the daily activities. The question is appropriate with the indicator “Writing” namely: Using language gambits.

**Question:**
Sinta : …, Ratna?
Ratna : She is a nurse.

a. Who is your mother  
   c. What does your mother do  

b. Where is your mother  
   d. What does your father do  

**Appropriate with indicator in English syllabus:**
The question asks about Ratna’s mother profession. The dialogue is very suitable to determine the indicator, namely: Using language gambits.

**Question:**
Andi : … ?
Pahrul : I play badminton twice a week.

a. How many sport events do you play  

b. How many times did you play badminton  

c. How many days did you play badminton  

d. How many times do you play badminton?

**Appropriate with indicator in English syllabus:**
This dialogue is used to know the fact about Andi’s badminton schedule. It is suitable with the indicator “Writing” namely: Using language gambits.

**Question:**
is – the – swimming – boy – the – in – river  
1 2 3 4 5 6 7  
The correct arrangement of the words above is…

a. 5 – 4 – 6 – 3 – 1 – 2 – 7  
b. 4 – 1 – 3 – 6 – 2 – 5 – 7  
c. 5 – 4 – 1 – 3 – 6 – 2 – 7  
d. 4 – 1 – 2 – 3 – 6 – 5 – 7

**Appropriate with indicator in English syllabus:**
This sentence is intended to know students’ ability in understanding Present Continuous Tense. The question’s form is suitable with the indicator, namely,
<table>
<thead>
<tr>
<th>Question:</th>
<th>The students… the school yard now.</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. are sweeping</td>
<td>c. to sweep</td>
</tr>
<tr>
<td>b. sweeps</td>
<td>d. swept</td>
</tr>
</tbody>
</table>

**Appropriate with indicator in English syllabus:**
The item is used to know students’ ability in grammatical. It is also suitable with the indicator: Denoting linguistic features in functional text.

<table>
<thead>
<tr>
<th>Question:</th>
<th>Rama : Are you reading a novel?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yahya : Yes, …</td>
<td>a. I do</td>
</tr>
<tr>
<td>b. I did</td>
<td>c. I am</td>
</tr>
<tr>
<td>d. You are</td>
<td></td>
</tr>
</tbody>
</table>

**Appropriate with indicator in English syllabus:**
The question also asks between the suitable to be and the subject. This item is appropriate to the indicator: Denoting linguistic features in functional text.

<table>
<thead>
<tr>
<th>Question:</th>
<th>A : Did you call me last night?</th>
</tr>
</thead>
<tbody>
<tr>
<td>B : No, I didn’t.</td>
<td>a. Are you alright?</td>
</tr>
<tr>
<td>A : …</td>
<td>b. Are you OK?</td>
</tr>
<tr>
<td>B : Yes, I’m sure.</td>
<td>c. Are you fine?</td>
</tr>
<tr>
<td>d. Are you sure?</td>
<td></td>
</tr>
</tbody>
</table>

**Appropriate with indicator in English syllabus:**
This item is intended to clarify the calling last night. It is really suitable with the indicator “Writing” namely: Using language gambits.

<table>
<thead>
<tr>
<th>Question:</th>
<th>Dede : Is it correct that an earthquake happened last night?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lila : …</td>
<td><strong>...</strong></td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>b. No, I didn’t.</td>
<td>d. Yes, I do.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Appropriate with indicator in English syllabus:**

The question is intended to clarify the correctness about an earthquake happened last night. The item is appropriate with the indicator, namely: Using language gambits.

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Question:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>She worked at Bank Danamon, …?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. doesn’t she</td>
<td>c. is she</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. didn’t she</td>
<td>d. did she</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Appropriate with indicator in English syllabus:**

The question asks about clarifying that is it true that she worked at Bank Danamon? The item used question tag to clarify. The item is suitable with the indicator: Determining suitable question tag in the sentence.

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Question:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is the color of Annajah building?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. White and Purple</td>
<td>b. Grey and Blue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Grey and White</td>
<td>d. Blue and Brown</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Inappropriate with indicator in English syllabus:**

The item asks about students’ ability in determining the word, color. So, Determining suitable vocabulary is not suitable with the indicator.

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Question:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adi : What fruits do you like?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laila : …</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. I like pizza.</td>
<td>d. I like apples and oranges.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Appropriate with indicator in English syllabus:**

The question asks about fruits which Adi like. This item is appropriate with the indicator “Writing” namely: Using language gambits.
<table>
<thead>
<tr>
<th>Question:</th>
<th>Mira  : Do you enjoy riding a bicycle to school?</th>
<th>Anna  : No, I don’t. It makes my legs are hurt.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mira  : I really enjoy it. It makes us health.</td>
<td>From the underlined expressions above, we can conclude that…</td>
</tr>
<tr>
<td></td>
<td>a. Anna really enjoys riding a bicycle to school.</td>
<td>b. Both Anna and Mira don’t enjoy riding a bicycle to school.</td>
</tr>
<tr>
<td></td>
<td>c. Both Anna and Mira enjoy riding a bicycle to school.</td>
<td>d. Mira really enjoy riding a bicycle to school.</td>
</tr>
</tbody>
</table>

**Appropriate with indicator in English syllabus:**
The dialogue discusses about the enjoyable to ride bicycle to school. That is suitable with the indicator “Writing” namely: Using language gambits.

<table>
<thead>
<tr>
<th>Question:</th>
<th>Agung : Would you like an ice cream, Din?</th>
<th>Andin : No, thanks. I don’t like it because it can make me fat.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The underlined word is an expression of…</td>
<td>a. Like b. Dislike c. Agree d. Disagree</td>
</tr>
</tbody>
</table>

**Appropriate with indicator in English syllabus:**
The dialogue discusses about the like of ice cream. That is suitable with the indicator “Writing” namely: Using language gambits.

<table>
<thead>
<tr>
<th>Question:</th>
<th>Rita  : Hi, friends! Tomorrow is a holiday. What shall we do?</th>
<th>Ratih : How about going to Carrefour Department store?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Erni  : … What about swimming?</td>
<td>Rita  : OK, so what time shall we go?</td>
</tr>
<tr>
<td></td>
<td>Erni  : Nine o’clock.</td>
<td>a. That’s not a good idea c. That’s fine</td>
</tr>
<tr>
<td></td>
<td>b. I agree with you</td>
<td>d. You’re right</td>
</tr>
</tbody>
</table>

**Appropriate with indicator in English syllabus:**
<table>
<thead>
<tr>
<th>Question:</th>
<th>This question asks about the different opinion in holiday plan. The item is appropriate with the indicator “Writing” namely: Using language gambits.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adi</td>
<td>Angelina, do you think a pilot is a good job?</td>
</tr>
<tr>
<td>Angelina</td>
<td>…</td>
</tr>
<tr>
<td>a. I think so</td>
<td>c. He flies a plane</td>
</tr>
<tr>
<td>b. Really?</td>
<td>d. He is handsome</td>
</tr>
<tr>
<td>24.</td>
<td></td>
</tr>
</tbody>
</table>

**Appropriate with indicator in English syllabus:**

The dialogue is intended to measure students’ ability in giving an opinion about pilot. This item is suitable with the indicator “Writing” namely: Using language gambits.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. do you think</td>
<td>c. do you agree</td>
</tr>
<tr>
<td>b. is the fact</td>
<td>d. does you think</td>
</tr>
<tr>
<td>25.</td>
<td></td>
</tr>
</tbody>
</table>

**Appropriate with indicator in English syllabus:**

The question asks about opinion of Harry Potter novel. So, this question is appropriate with the indicator “Writing” namely: Using language gambits.

<table>
<thead>
<tr>
<th>Question:</th>
<th>The vehicles which there are in the village, except?</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Pedicab</td>
<td>c. Bicycle</td>
</tr>
<tr>
<td>b. Cart</td>
<td>d. Plane</td>
</tr>
<tr>
<td>26.</td>
<td></td>
</tr>
</tbody>
</table>

**Inappropriate with indicator in English syllabus:**

This question talks about kinds of vehicles in the village. The indicator Determining suitable vocabulary is not suitable with the syllabus.

<table>
<thead>
<tr>
<th>Question:</th>
<th>Mrs. Inayah is going to a supermarket. Now she is talking to one of the shop assistants.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shop assistant : Can … , Madam?</td>
<td>Mrs. Inayah : Well, I need a shirt for my husband.</td>
</tr>
<tr>
<td>Question</td>
<td>Options</td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>Shop assistant : We have many kinds of shirt. Which shirt do you want?</td>
<td>a. I help you – large c. do it – large b. help you – small d. you help me – small</td>
</tr>
<tr>
<td>Mrs. Inayah : He is tall and fat.</td>
<td></td>
</tr>
<tr>
<td>Shop assistant : You had better take this... one.</td>
<td></td>
</tr>
<tr>
<td>Tasya : I don’t have any pen. ..., please?</td>
<td>a. What is this c. May I have are b. Take me the pen d. Give me the pen</td>
</tr>
<tr>
<td>Rudi : Sure. Here you are.</td>
<td></td>
</tr>
<tr>
<td>Dany : May I borrow your book, please?</td>
<td>a. Thank you c. I’m sorry. I’m using it now b. Yes, I do d. What?</td>
</tr>
<tr>
<td>Yuni : ...</td>
<td></td>
</tr>
<tr>
<td>... he recovered from his illness, he had to take a rest.</td>
<td>a. After c. Then</td>
</tr>
</tbody>
</table>
### Question:

**Preparing the printer**

Follow these steps to prepare your printer.

1. After that, remove the tapes from the printer as shown.
2. Finally, to remove the Protective Material on the right side.
3. Next, open the Paper Support and the Front Cover.
4. First, place the printer on a flat surface.

Rearrange the sentences into a good paragraph…

<table>
<thead>
<tr>
<th>a. 1 – 2 – 3 – 4</th>
<th>b. 4 – 2 – 1 – 3</th>
<th>c. 4 – 1 – 3 – 2</th>
<th>d. 1 – 4 – 3 – 2</th>
</tr>
</thead>
</table>

### Appropriate with indicator in English syllabus:

This item is intended to know students’ ability in understanding Procedural Text. The question’s form is suitable with the indicator, namely, Jumbling text/sentence/word into good arrangement.

### Question:

Eko : Would you like a glass of ice tea?
Huda : …
Eko : All right, I will take it for you.

<table>
<thead>
<tr>
<th>a. Yes, please. I am thirsty.</th>
<th>c. No, thanks. I am not thirsty.</th>
</tr>
</thead>
<tbody>
<tr>
<td>b. I am sorry. I am very thirsty.</td>
<td>d. I’d love to. But I am not thirsty.</td>
</tr>
</tbody>
</table>

### Appropriate with indicator in English syllabus:

The question is intended students’ ability in understanding the suitable answer from the expression how to offering something. It is very suitable with the indicator “Writing” namely: Using language gambits.
**Mother:** Can you get me some tomatoes from the refrigerator?  
**Ricky:** …

b. Yes, that’s right. d. No, thanks.

**Appropriate with indicator in English syllabus:**  
The question is intended students’ ability in understanding the suitable answer from the expression how to offering something. It is very suitable with the indicator “Writing” namely: Using language gambits.

**Question:**  
What kind of text is the text above?

**42. Inappropriate with indicator in English syllabus:**  
This item is intended to know the kinds of text. So, Identifying the kinds of text that are meant by the test is not suitable with the indicator.

**Question:**  
Shop Assistant: Wow! You bought so many things today, Mrs. Kris. Would you like me to carry them to your car?  
Mrs. Kris: Oh! Thank you. That’s very kind of you.  
Shop Assistant: No problem.

**Which expressions show Asking and Offering a Favor?**

**Appropriate with indicator in English syllabus:**  
The item is intended to measure students’ ability in identifying the expression of asking and offering a favor. The question is appropriate with the indicator “Writing” namely: Using language gambits.

**Question:**  
Make 2 sentences with using Present Continuous Tense!

**44. Appropriate with indicator in English syllabus:**  
The item is used to know students’ ability in grammatical. It is also suitable with the indicator: Denoting linguistic features in functional text.
<table>
<thead>
<tr>
<th>Question:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Give explanation about:</td>
</tr>
<tr>
<td>a. What is the purpose of the Procedural Text?</td>
</tr>
<tr>
<td>b. What is the general structure of the text?</td>
</tr>
</tbody>
</table>

### Appropriate with indicator in English syllabus:

This item include 2 indicators: Identifying communicative purpose and rhetorical steps in descriptive/procedure text. So, it is appropriate to the indicators.

Based on the tables above and the explanations in the description of the data, we know that some indicators are not entered to the syllabus but they are included in the test. So the test becomes inappropriate. Actually, the inappropriate indicators are still included in the competence standard and basic competency. In the other word, there are some indicators that are not included in the test whereas the indicators mention in the syllabus. The test will be better if all the indicators in the syllabus are entered in the test.

Beside that, there are also some of expansion items number which are not proportional to the test. There are indicators which has many items number. On the other hand, there are some indicators which has few items number. Indeed they just have one item number.

The possibilities reason why the materials are more appeared in the items of the test are the materials in the indicators are easy to learn by the students and also very suitable with the real context in daily activities. On the other word, the possibilities reason why the materials are less appeared in the items of the test are the materials in the indicators are difficult to learn by the students, the time is very limited, and the materials have high difficulties.

The effects, that arise from the problems, are the students do not master all the suggested materials in the curriculum and the goals of the curriculum do not achieved well.

3. **The Interpretation of the Data**

The following table describes the total frequency of the appropriateness and the inappropriateness of English Summative test items to the syllabus based on the
data of item analysis result above. Based on the data of items analysis result, we can
see the table as follow:

Table 3.6
The Analysis Result of Appropriateness and Inappropriateness of English
Summative Test Items at Even Semester of the First Grade Students of MTs.
Annajah Petukangan

<table>
<thead>
<tr>
<th>No</th>
<th>Summative Test Items Area</th>
<th>Items Number</th>
<th>Total Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The appropriateness test items based on indicators in English syllabus</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 30, 31, 32, 33, 34, 35, 39, 40, 41, 43, 44, 45</td>
<td>38 Items = 84 %</td>
</tr>
<tr>
<td>2.</td>
<td>The inappropriateness test items based on indicators in English syllabus</td>
<td>13, 16, 26, 36, 37, 38, 42</td>
<td>7 Items = 16 %</td>
</tr>
</tbody>
</table>

Table 3.6 shows that the appropriateness based on the syllabus is about 84 % consist of 38 items. This percentage obviously falls into the level of 76 – 100 % which means good. Then, the inappropriateness items based on syllabus are only 16 % which consist of 7 items. The calculations are mention below.

Based on the data of the appropriateness items number, the writer analyzes:

\[
P = \frac{f}{n} \times 100 \%
\]

Note:
\[
P = \frac{f}{n} \times 100 \%
\]

Based on indicators in the English syllabus

\[
P = \frac{38}{45} \times 100 \%
\]

= 84%

Based on the data of the inappropriateness items number, the writer analyzes:

\[
P = \frac{f}{n} \times 100 \%
\]

Note:

P = Percentage
f = Frequency
n = Number / Amount

Based on indicators in the English syllabus

\[
P = \frac{7}{45} \times 100 \%
\]

= 16%
According to the item analysis result above, the writer concludes that the English summative test which is administrated in the first grade students of MTs Annajah has good validity.

From the interview with the English teacher as the test designer, she said that “she constructed the test individually not by the teacher’s group work. Before she made the test, firstly she made the blue print of the questions.”

She made the test based on the indicators in the English syllabus of the even semesters. According to her, “not all the items in the indicator have to be entered in the test. The important thing is we may not come out from the competence standard and basic competency. So, its okay, if what the curriculum asked to do, we have done so far.”

Therefore, it can be understood that a test designer have already understood how to make a good test. They recognized that the summative test must be in line with the recommended curriculum and also with the English syllabus.
CHAPTER IV
CONCLUSIONS AND SUGGESTIONS

A. Conclusions

Based on the data in the previous chapter, the writer would like to elaborate the appropriateness of English summative test of the even semester for the first grade students of MTs Annajah Petukangan to the English syllabus used. This summative test consists of 45 items (40 items of multiple choice and 5 items of essay).

The item test that measures the suggested indicators in the syllabus is 38 items and the inappropriateness item is 7 items. It means that the appropriateness items cover 84% from the indicator and the inappropriateness items is 16%.

The appropriateness percentage obviously falls in to the level of 76 – 100% which means good. The test contains a representative sample of material and the learning objectives of even semester, it can be seen from the table that many learning objectives, and language gambits are included in the test. Therefore, it can be understood that the English summative test items at even semester on the first grade students of MTs Annajah has good content validity.

Beside that, if we want to construct a test, we have to pay attention to the criterion of a good test. One of the criterions is validity, especially for this time is content validity. Based on this research, although there are some items which are not appropriate with the indicators in the syllabus, but that items do not come out from the Competence Standard and Basic Competency in the English curriculum. And then, all of the indicators, which are made by the school, should be better if they are entered as the items in the English test.

B. Suggestions
The writer would like to give some suggestions addressed to the test designers as feedback of the research result:

1. The test designers should have a good preparation before making a test, by considering the principles of constructing good test items.
2. The test designers should consider that the test items should be appropriate to the syllabus category used; the items that do not fulfill to the curriculum should be discarded.
3. The items that do not fulfill to the syllabus should be revised so that they can be used for the next evaluation.

She also would like to give some suggestions addressed to the school as follow:

1. The school must give chances to the English teachers to follow trainings of constructing a good test.
2. The school must supervise the teachers’ work in constructing the test, whether the constructed test has fulfilled characteristics of good test or not.
REFERENCES


