AN ANALYSIS OF CONTENT VALIDITY OF ENGLISH SUMMATIVE TEST IN THE ODD SEMESTER ACADEMIC YEAR 2008/2009

(A Case Study at the First Year of Madrasah Tsanawiyah Pembangunan UIN, Jakarta)

A “Skripsi”
Presented to the Faculty of Tarbiyah and Teacher’s Training
In Partial Fulfillment of the Requirements
for the degree of S.Pd (Bachelor of Art) in English Language Teaching

by
AGUSTIN FAJAR TRIANINGSIH
204014003200

Approved By The Advisor

Dr. M. Farkhan, M. Pd.

DEPARTMENT OF ENGLISH EDUCATION
FACULTY OF TARBIYAH AND TEACHERS TRAINING
SYARIF HIDAYATULLAH STATE ISLAMIC UNIVERSITY
JAKARTA
2010
AN ANALYSIS OF CONTENT VALIDITY OF ENGLISH
SUMMATIVE TEST IN THE ODD SEMESTER
ACADEMIC YEAR 2008/2009

(A Case Study at the First Year of Madrasah Tsanawiyah Pembangunan UIN, Jakarta)

A “Skripsi”
Presented to the Faculty of Tarbiyah and Teacher’s Training
In Partial Fulfillment of the Requirements
for the degree of S.Pd (Bachelor of Art) in English Language Teaching

by
AGUSTIN FAJAR TRIANINGSIH
204014003200

DEPARTMENT OF ENGLISH EDUCATION
FACULTY OF TARBIIYAH AND TEACHERS TRAINING
SYARIF HIDAYATULLAH STATE ISLAMIC UNIVERSITY
JAKARTA
2010
ENDORSEMENT SHEET

The examination committee the Faculty of Tarbiyah and Teacher’s Training certifies that the “Skripsi” (Scientific Paper) entitled “AN ANALYSIS OF CONTENT VALIDITY OF ENGLISH SUMMATIVE TEST IN THE ODD SEMESTER ACADEMIC YEAR 2008/2009 Conform to the English Syllabus” (A Case Study at the First Year of Madrasah Tsanawiyah Pembangunan UIN, Jakarta) written by Agustin Fajat Trianingsih, whose student’s registration number: 2040 1400 3200 was examined by the committee on March 19th 2010 and was declared to have passed, and therefore, fulfilled one of the requirements for the academic title of “S.Pd.” (Bachelor of Art) in English Language Education at the Department of English Education.

Jakarta, March 2010
Examination Committee

Chairman : Drs. Syauki, M.Pd (____________)
19641212 199103 1 002

Secretary : Neneng Sunengsih, S.Pd (____________)
19730625 199903 2 001

Examiner I : Drs. Syauki, M.Pd (____________)
19641212 199103 1 002

Examiner II : Drs. Zaenal Arifin Toy, M. Sc. (____________)
150 031 215

Acknowledgement by,
Dean of Tarbiyah and Teacher’s Training Education

Prof. Dr. Dede Rosyada, M. A.
19571005 198703 1 003
CHAPTER ONE
INTRODUCTION

A. Background of Study

English has become one of the subjects to be taught to all levels of educational Institution in Indonesia, starting from primary school as a local-content subject up to university level as a complimentary subject. Every level has evaluation to give an information how successful in teaching learning activity. The information gained through this evaluation will be very useful to make improvement in the future.

Evaluation is an activity of getting information to be used to make educational decision. As the 2004 English Curriculum said that “Penilaian menekankan pada proses dan hasil belajar dalam upaya penguasaan atau pencapaian suatu kompetensi” (evaluation emphasized on the student process and learning outcomes to mastery a certain competence.¹ The writer needs to know periodically how successful their efforts have been, so that they can decide which practice to continue and which to change. Sometimes the decisions are not about the students, but more often they are about the instructional.

Evaluation is a systematic process of collecting, analyzing and interpreting information to determine the extent to which pupils are achieving instructional objectives. Evaluation in teaching learning English is important because it is used to know the appropriateness and attainability of the instructional objective, the usefulness of the instructional materials and the effectiveness of the instructional methods.

Evaluation is continuous process which underlies all good teaching and learning. Carefully collected evaluation data help the teachers understand the learners, plan learning experience for them and determine the extent to which the instructional objective are being achieve. Evaluation also supplies the information on which to base decisions concerning curriculum development. The result of evaluation will be very helpful to curriculum designer in improving and developing organizational arrangements of the curriculum.

There are many methods for collecting information for evaluation purpose. One of them is by using a test. Tes adalah suatu alat atau prosedur yang sistematis dan objektif untuk memperoleh data-data atau keterangan-keterangan yang diinginkan tentang seseorang, dengan cara yang boleh dikatakan tepat dan cepat (Test is a systematic and objective procedure to obtain the data or information about learner by an appropriate technique). There are two kinds of test that are usually used at the school, they are formative and summative test. Formative test is usually made by the teacher at the end of a unit in the course book. Summative test is usually made by a team, given at the end of the schools year and it is held in every school together in the same time.

Test is a process or device in education to get the information of the students achievements in order achieve the learning objective. A test should be constructed well to measure the student achievement, so that it is able to distinguish between students who studied well and they have not. Sometimes

---

students usually tend to study harder when they have examination and they only study the material that they expect to be tested. It is not easy job for the teacher to make a good test because there are some characteristic and requirements, which must be involved and fulfilled.

The characteristics of a good test, which must considered by a test maker or teacher are, validity, reliability and practicality. The most important quality to consider when selecting or constructing an evaluation instrument is validity. They only evaluate certain aspects from the whole aspects as suggested by the curriculum.

Therefore, all summative tests should be tested based on the materials as covered by the curriculum. If not, they can not be considered as a good test. They, of course, fail to get the accurate data about the student’s achievement. To fulfill these criteria the test makers should take school based on curriculum (KTSP) into their consideration. The curriculum is used based on National Education Permendiknas No. 22 tahun 2006. About this curriculum, Mulyasa said:

"Pengembangan standar kompetensi dan kompetensi dasar kedalam kurikulum operasional Tingkat Satuan Pendidikan, merupakan tanggung jawab satuan pendidikan masing-masing. Oleh karena itu, sebutan untuk kurikulum ini adalah KTSP, singkatan dari Kurikulum Tingkat Satuan Pendidikan.”

As one of Islamic Junior High Schools, Madrasah Tsanawiyah Pembangunan UIN has applied the concept of the curriculum 2006 in teaching and learning English; and evaluating the students’ progress.

Institutionally, the school has already constructed English summative test fulfilling the criteria of validity. It has already made the curriculum as the basis for preparing the tests’ items. However, it is still a big question whether the English summative test really reflect the curriculum or not. It, of course, affects the test’s validity.

---

5E. Mulyasa, *Kurikulum Tingkat Satuan Pendidikan*, (Jakarta: Remaja Rosdakarya, 2007), p. 11-12
Because of the importance of the test’s validity, the writer would like to write the paper entitled “An Analysis of Content Validity of English Summative Test in the Odd Semester Academic Year 2008/2009” (A case study at the first Grade of Madrasah Tsanawiyah Pembangunan UIN, Jakarta).

B. Statement of Problem

Based upon the background of the study, the statement of the problem in this study is as the follows “To what extent does the English summative test for the first grade learners conform to the English Syllabus of Madrasah Tsanawitah Pembangunan UIN, Jakarta?”

C. Limitation and Scope of the Study

The writer limits this paper on the analysis on test items of English summative test and English Syllabus in the first year students of Madrasah Tsanawiyah Pembangunan UIN, Jakarta.

D. Objective of the Study

The Objective of the study is to find out if the test items of English Summative test for the first year students of Madrasah Tsanawiyah Pembangunan have content validity accordance to the English Syllabus. Through this study, it is also hoped that the teacher can get clear description about the quality of the test items of English Summative test.
E. Method of Study

This study employs a descriptive comparative method by which the writer compares the items of English summative test with the syllabus used by Madrasah Tsanawiyah Pembangunan UIN, Jakarta. The items will be analytically compared to the important components of the curriculum to know their degree of representativeness.
- Artha Gading
- 23 mei 2010 at Monas
- At Kebumen
- At Empu Tantular
Tour Lampung with Hijau

Daun

Jakarta, 23 April 2010
Road Manager

UAN
A. Syllabus

1. Definition of Syllabus

The discussion of education is absolutely related to the curriculum and syllabus. Most people in education field, at this point teachers or educators, are familiar with these terms. To more understand about the syllabus, there are many definitions about the syllabus from the experts. According to Tarigan “Silabus adalah suatu pemerian mengenai aneka isi atau bobot satu kursus pengajaran dan urutan atau susunan bahan yang akan diajarkan.” (Syllabus is the list of various materials that should be taught by the teacher in a course or school for the learner).\(^1\) Another statement come from Penny Ur, syllabus is a reference of the course a teacher teaches or intents to teach.\(^2\)

For more detail Dr. E. Mulyasa stated that:

Dalam KTSP, syllabus merupakan penjabaran standar kompetensi dan kompetensi dasar kedalam materi pembelajaran, kegiatan pembelajaran dan

\(^1\) Henry Guntur Tarigan, *Pengajaran Kompetensi Bahasa*, (Bandung: Angkasa, 1984), p.78
indicator pencapaian keterampilan untuk penilaian hasil belajar. ” (In KTSP, Syllabus is spelling out standard competence and basic competence into learning material, learning activity and indicator to evaluate the student progress).”

While Robertson as quoted by Tarigan defined that syllabus is a statement about planning of each part of curriculum expecting the element of evaluation itself. Moreover, Mary Finacchiaro and Christopher Brumfit said that a syllabus generally provides a listing of content to be learned but does not suggest methods activities and measures of evaluation.

Thus, it can be summarized that syllabus is a list of several materials that should be learned by the students in order to fulfill the demand of curriculum that is developed continuously.

### 2. The Purpose of Syllabus

According to Howard B. Altman, the primary of the syllabus is to communicate to one’s students what the course is about, why the course is taught, where it going and what will be required of the students for them to complete the course with a passing grade.4

According diamond R.M the purpose of the syllabus is:

- **Syllabus as a contract**, the syllabus serves to set forth what is expected during the term of the contract and, typically a semester and to guide the behaviors of both parties.

- **Syllabus as a permanent record** by providing details of what covered, what students were expect to do and hoe these outcomes and performances are assessed.

---

3 E. Mulyasa, *Kurikulum Tingkat Satuan Pendidikan*, (Jakarta: Remaja Rosdakarya, 2007), p. 190

c. **Syllabus as a learning tool**, a course syllabus can serve as a highly effective facilitator of student learning.\(^5\)

Based on the explanation above, the writer can conclude that the main purpose of the syllabus is to communicate to the students what the course is about, it presumes that we have some idea about what we think the course should accomplish. It requires that we have planned.

### 3. The Component of Syllabus

The function of the syllabus is to guide the teaching and learning process. It contains the learning objectives, material, indicator, sources and the allocation of time. According to Jo McDonough, the broad syllabus outline will in turn have direct implications for the more detailed design selection of materials and testing, the planning of individual lesson, and the management of the classroom itself.\(^6\)

In order to arrange the syllabus, several components are necessary. Mulyasa said that there is minimally five components of syllabus:

   a. Kompetensi Dasar
   b. Standar Kompetensi
   c. Indikator
   d. Materi Standar
   e. Standar Process
   f. Standar Penilaian\(^7\)

### 4. Format of The Syllabus

A syllabus is designed based on school program and students’ degree. In English syllabus 2006 at least six components, there are Standard competence, Basic competence, indicator, material, learning activity, evaluation, the allocation of time and resources.

---


\(^7\) E. Mulyasa, *Op cit*, p. 191
According to the English syllabus that used in Madrasah Tsanawiyah Pembangunan UIN Jakarta covers four standard competences. There are reading, speaking, listening and writing.

In addition, reading and writing are appears in the English summative test because their concerns mastery of how to understand on the written task. In the other hand, listening and speaking are not.

It is obvious that only reading and writing focused on this research because they should be conformed by the English summative test. The writer expects the composer of:

| Table 2.1 |
| The Learning Objective of the English Syllabus |

<table>
<thead>
<tr>
<th>The learning objective of English syllabus</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Reading</strong></td>
</tr>
<tr>
<td>a. Find the ideas of reading text.</td>
</tr>
<tr>
<td>b. Find the explicit topic of reading text.</td>
</tr>
<tr>
<td>c. Find the implicit topic of reading text.</td>
</tr>
<tr>
<td>d. Find the detail information.</td>
</tr>
<tr>
<td>e. Find the explicit supporting details.</td>
</tr>
<tr>
<td>f. Find the implicit supporting details.</td>
</tr>
<tr>
<td>g. Interpret the meaning of word, phrase and sentence based on the context.</td>
</tr>
<tr>
<td>h. Get the happiness.</td>
</tr>
<tr>
<td><strong>2. Writing</strong></td>
</tr>
<tr>
<td>a. Change the simple sentence with the correct spelling a correct punctuation.</td>
</tr>
<tr>
<td>b. Write the simple sentences which dictated by teacher.</td>
</tr>
<tr>
<td>c. Complete the simple dialogue or short paragraph (narration or description)</td>
</tr>
</tbody>
</table>
### Table 2.2

The Topics and Subtopics of the English Syllabus

<table>
<thead>
<tr>
<th>Topics</th>
<th>Sub topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self identity</td>
<td>A. Introduction</td>
</tr>
<tr>
<td></td>
<td>B. Personal Appearance</td>
</tr>
<tr>
<td>School life</td>
<td>A. At the School</td>
</tr>
<tr>
<td></td>
<td>B. In the classroom</td>
</tr>
<tr>
<td></td>
<td>C. in the library</td>
</tr>
<tr>
<td></td>
<td>D. In the canteen</td>
</tr>
<tr>
<td></td>
<td>E. In the laboratory</td>
</tr>
<tr>
<td></td>
<td>F. School garden</td>
</tr>
<tr>
<td></td>
<td>G. School activity</td>
</tr>
<tr>
<td></td>
<td>H. School rules</td>
</tr>
<tr>
<td></td>
<td>I. Scout</td>
</tr>
<tr>
<td>Family life</td>
<td>A. Family members</td>
</tr>
<tr>
<td></td>
<td>B. House and tools</td>
</tr>
<tr>
<td></td>
<td>C. Family activities</td>
</tr>
<tr>
<td></td>
<td>D. Daily activities</td>
</tr>
<tr>
<td></td>
<td>E. Neighborhoods</td>
</tr>
<tr>
<td></td>
<td>F. Pet</td>
</tr>
<tr>
<td></td>
<td>G. Wealth</td>
</tr>
<tr>
<td>Profession</td>
<td>A. In the work place</td>
</tr>
<tr>
<td></td>
<td>B. Kinds of work</td>
</tr>
<tr>
<td></td>
<td>C. Situation of work</td>
</tr>
<tr>
<td>No</td>
<td>English functional skill of the English syllabus</td>
</tr>
<tr>
<td>----</td>
<td>------------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td><strong>Comprehend and Express</strong></td>
</tr>
<tr>
<td>1</td>
<td>Self introduction</td>
</tr>
<tr>
<td>2</td>
<td>Introduction someone to the others</td>
</tr>
<tr>
<td>3</td>
<td>Greeting in the first meeting/ goodbye</td>
</tr>
<tr>
<td>4</td>
<td>Thanks</td>
</tr>
<tr>
<td>5</td>
<td>Like/ dislike</td>
</tr>
<tr>
<td>6</td>
<td>Apology</td>
</tr>
<tr>
<td>7</td>
<td>Imperative</td>
</tr>
<tr>
<td>8</td>
<td>Factual information (identifying, asking, explaining and telling)</td>
</tr>
<tr>
<td>9</td>
<td>Want</td>
</tr>
<tr>
<td>10</td>
<td>Capability / incapability</td>
</tr>
<tr>
<td>11</td>
<td>Possessives</td>
</tr>
<tr>
<td>12</td>
<td>Profession</td>
</tr>
<tr>
<td>13</td>
<td>Choice</td>
</tr>
<tr>
<td>14</td>
<td>Adverb of place</td>
</tr>
<tr>
<td>15</td>
<td>Existence someone/ something</td>
</tr>
<tr>
<td>16</td>
<td>Explanation about someone, object, &amp; event which its amount more than one</td>
</tr>
<tr>
<td>17</td>
<td>Quantity of uncountable noun</td>
</tr>
<tr>
<td>18</td>
<td>Explanation about something which taking place</td>
</tr>
<tr>
<td>19</td>
<td>Habitation</td>
</tr>
<tr>
<td>20</td>
<td>Past event</td>
</tr>
<tr>
<td>21</td>
<td>Future event</td>
</tr>
<tr>
<td>22</td>
<td>Amount something</td>
</tr>
<tr>
<td>23</td>
<td>Obligation</td>
</tr>
</tbody>
</table>
B. Test

1. Definition of Test

Michael T. Nietzel says that a test is a systematic procedure for observing and describing a person’s behavior in a standard situation.\(^8\) To understand more about test, there are many definitions about the test from the experts. According to Prof. Drs. Anas Sudijono “Tes adalah alat atau prosedur yang dipergunakan dalam rangka pengukuran dan penilaian.” (Test is a tool or procedure used for measuring and scoring).\(^9\) Other statement came from Drs. M. Chabib Thoha, M.A compiled from Encyclopedia of Educational Evaluation “Test is any series of questions or exercise or other means of measuring the skills, intelligence, capacities of attitudes or an individual or group.”\(^10\)

For more detail Wayan Murkancana and PPN Sumartana stated that:

“Tes adalah suatu cara untuk mengadakan penelitian yang berbentuk suatu tugas yang harus oleh anak atau sekelompok anak sehingga menghasilkan suatu nilai tentang tingkah laku atau prestasi anak dengan nilai standar yang ditetapkan.” (A test is a technique of measurement which can be a task or correlated task which must be done by a learner or a group of learners, in order to yield a norm about his or their achievement and can be compared with score achieved by a learner or a group of learners, or with a standard norm).\(^11\)

While L.J Cronbach in his book *Essential Psychological Testing*, states that test is a systematic procedure for observing a person’s behavior and describing it with the aid of either a numerical scale or a category system.\(^12\)

---


Usually such scale or category system provides scores which must be interpreted. Jum. C. Nunally says that a test is standardized situation that provides an individual with a numerical score.\textsuperscript{13} Some of the most essential ingredients of standardization are:

a. All students should answer the same questions.

b. Instruction should be clear and the same instruction should be given to all students.

c. No student should be given any advantage not given to all students.

d. A predetermined system of scoring should uniformly be applied to the answer of all students.\textsuperscript{14}

Based on the above opinion, the writer can conclude that a test is a tool or device either in form of question, order or direction used to measure and value the student’s achievement in mastering the material given by the teacher in accordance to the curriculum.

2. Purpose of the Test

According to Drs. M. Ngalim Purwanto, “Tes dipergunakan untuk menilai hasil belajar pelajaran yang diberikan guru kepada murid-muridnya atau dosen kepada mahasiswa (Test is used to evaluate the achievement of the material given by teacher to students, or lecture to students)”.\textsuperscript{15}

According to Anthony J. Nitko the purpose of test is:

a. Helpful to obtain measures the instructional process as well as the products of instruction.

b. Suggests various kinds of instructional process constructs and variables and how they might be measured.


\textsuperscript{14}Ibid.

\textsuperscript{15}M. Ngalim Purwanto, \textit{Prinsip-Prinsip dan Teknik Evaluasi Pendidikan} (Bandung: PT Remaja Rosda Karya), p.33
c. To help in assigning grades to students.
d. Providing feedback to students about their learning.
e. Provide feedback to the teacher as well as to students’ information about how well students have learned and how well the teacher has taught.
f. Serve as motivators for students to study.\(^\text{16}\)

While, according to Herbert and William the purpose of test are to:

a. Facilitate students’ progress.
b. Assess student progress
c. Understand the individual student.
d. Facilitate self understanding by the student.
e. Evaluate instructional programs.
f. Assist in administrative judgments.
g. Contribute to knowledge of abilities and instruction.\(^\text{17}\)

Based on the explanation above the writer can conclude that the main purpose of the test is to know whether the curriculum have been achieve or not, to know the student’s effort and to know how successfully teaching learning activity.

3. Types of Test and Types of Test Item

a. Types of Test.

According to Petunjuk Pelaksanaan Kurikulum 2006, there are three types of evaluation. They are:

1). **Formative Test** is given at the end of the lesson unit, functioning as feedback to know the degree of the students’ achievement individually or as group or to see how far the students have reached the instructional objectives. Then, the results of the test will be used to improve the implementation of teaching-learning process.


2). **Summative Test** is given at the end of semester. It has to cover all the themes and sub themes.

3). **Nationally Examination** is given at the end of third year and test is held nationally (UN).

Benjamin Bloom has developed the concept of formative and summative evaluation as follows:18

1). **Formative Test** is given during the course of instruction (at the end of lesson unit); its purpose is how to show which aspects of the chapter the student has mastered and where remedial work is necessary. The formative test is usually graded on a pass-fail basis, and students who fail are given the opportunity to study and then taken the test again. Generally, formative test at the public school is usually called daily review, as stated by Prof Dr Suharsimi Arikunto “Dalam pelaksanaannya disekolah tes formative merupakan ulangan harian”.19

2). **Summative Test** is usually given at the end of a marking period and measures the “sum” total of the material covered. On this type of a test, students are usually ranked and graded. The summative test covers all the subject unit (themes and topics) and measures the general instructional objective, as stated by Wilmar Tinambunan, “…the test must reflect the content of the whole course and the test must be able to

---


determine the extent to which the instructional objectives have been achieved by pupils”.

b. Types of Test Item

The question, exercise and task appearing on a test are called items. The kinds of items on test are:

1) **The latter type of items** called choice items, which include multiple choice items, true false and matching.

a) Multiple Choices

The multiple choices consist of the stem and a set of alternatives. The stem is presented as a question or incomplete statement which the students answer or complete by selecting one of several alternatives.

According to Jum C. Nunally, “By far the most popular types of objective item is that in which the student is required to choose one alternative response to a problem or question”.

While J. Stanley Ahmann and Marvin D. Glock: “A multiple choice item is one of which a direct question or incomplete statement is presented and a number of possible response or option are given.” The multiple-choice item consists of a problem and a list of suggested solution. The problem may be stated in the form of a direct question or and incomplete and it’s called the stem of the item.

---

Meanwhile Julian C Stanley states that “A multiple choice test is made up of items each of which presents two or more responses, only one of which is correct and definitely better than others.”

Constructing a test item is not easy, and neither is constructing a multiple choice item test. The multiple choice test consists of two main points, the option and the stem. In constructing the stem, there are several rules that should be fulfilled to make a good stem:

(1). The stem of the item should be meaningful by itself and should present a definite problem.
(2). The item stem should include as much of the item as possible and should be free of irrelevant material.
(3). Use negatively stated item only when significant learning outcomes require it.
(4). Verbal association between the stem and the correct answer should be avoided.

Same as the stem, there are also some rules in constructing the option to make it good. The rules are:

(1). All of alternatives should be grammatically consistent with the stem of the item.
(2). The relative length of the alternatives should not provide a clue to answer.
(3). The correct answer should appear in each of the alternatives positions approximately an equal number of times, but in random order.
(4). Use the special alternatives such as ‘none of the above’ or ‘all of the above’ sparingly.

And the most important rules for the construction of multiple-choice items are as follows:

(1). The problem should clearly point to the theme of correct alternative answer.

---

26 Ibid.
(2). Incorrect alternatives should be plausibly related to the problem.
(3). Correct alternatives should not be consistently different in appearance from incorrect alternatives.
(4). Alternatives should be randomly ordered for each item.
(5). Avoid irrelevant sources of difficulty in the statement of the problem or in the alternatives.
(6). Avoid including material in the problem, which is unrelated to the theme of the intended response.
(7). Do not employ alternatives which say “none of the above”, “both a and c above”, “all of the above”, etc.
(8). Avoid grammatical cues and sentences structures that give away the correct alternative.
(9). Use negative sparingly in problem statement.
(10). Ensure that item content relates to important aspects of the subject matter.  

The advantages of Multiple-choice items include their:

(1). The multiple-choice item is adaptable to subject matter content areas as well as different levels of behavior.
(2). The structure of a premise with four or five alternatives provides less chance for guessing the correct response than the true-false item does.
(3). One advantage of the multiple-choice item over the true-false item is that pupils cannot receive for simply knowing that a statement is incorrect; they must also know what is correct.
(4). Four or five option of each multiple-choice test provide more incorrect choice for selections of responses by the student who does not know the best or correct answer.
(5). The difficulty of each multiple-choice item can be controlled by changing the alternatives.
(6). Multiple-choice items are amenable to item analysis which enables the teacher to determine how well the items functioned with the students tested and how well each alternative functioned I discriminating between the higher achieving and lower achieving students.  

28 Wilmar Tinambunan, Op. Cit, p. 75-76
Meanwhile, the disadvantages of Multiple-choice items are:

(1). The technique test only recognition knowledge.
(2). Guessing may have a considerable but unknowable effect on test scores.
(3). The technique severely restricts what can be tested.
(4). It is difficult to write successful item.
(5). Backwash may be harmful.
(6). Cheating may be facilitated.29

b) True-False

According to Victor H. Noll, true-false item usually consist of a declarative sentence, to which that examinee responds by marking it true or false.30 Other statement came from Jum C. Nunally, true-false is presented with a statement to be marking as either true or false.31

True-False test is a test that requires the students to determine whether the statement is true or false. This response is depending on the instruction, by marking the initial of true (T) or false (F).

The common uses of true or false items are:

(1). To measure the ability to identify the correctness of statements of facts, definition of terms and statements of principles.
(2). To measure the pupil’s ability to distinguish fact from opinion.
(3). To measure the aspect of understanding.
(4). To measure the simple aspect of logic as illustrated in the example.32

---

According Robert L. Thorndike and Elizabeth Hagen, there are many things concerning true-false items, they are:

(1). Beware of negative statements and particularly of double negatives.
   Example:
   \[
   \begin{array}{ll}
   T & F \\
   \end{array}
   \]
   The constitution does not provide that no state law can deny a citizen the right to vote.

(2). Beware of items that include more than one idea in the statement, especially if one is true and another is false.
   Example:
   \[
   \begin{array}{ll}
   T & F \\
   \end{array}
   \]
   The president has the power to make treaties with foreign countries, but the senate must approve them by a majority of votes.

(3). Beware of items where the correct answer depends upon one insignificant word, phrase or letter.
   \[
   \begin{array}{ll}
   T & F \\
   \end{array}
   \]
   Ulysses Grant’s middle name was Sampson

(4). Beware of giving cues to the correct answer by the length of the item.\(^{33}\)

c) Matching

According to J. Stanley Ahmann, the matching test in its simplest form consists of two list of items and a set of instruction for matching each of the items in the first with one in the second.\(^{34}\)

While Wilmar Tinambunan states that the matching items are useful in measuring student’s ability to make association, interpretation or measures knowledge of a series of fact. Beside that, the matching items can be used for a


\(^{34}\) J. Stanley Ahmann & Marvin D. Glock, Op.Cit., p.96
large quantity of associated factual material to be measured in a small amount of space while the students’ time needed to respond is a relatively short.\textsuperscript{35}

Matching test is a test that requires the students to match the question with its answer. The student should find and match the statement with its answer. For instance:\textsuperscript{36}

Match the sentences with its classification:

<table>
<thead>
<tr>
<th>The answer</th>
<th>Column A</th>
<th>Column B</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>1. During the winter days are short and the nights are long.</td>
<td>A. Simple sentence</td>
</tr>
<tr>
<td>A</td>
<td>2. Jane rode to school on her bicycle.</td>
<td>B. Complex sentences</td>
</tr>
<tr>
<td>B</td>
<td>3. Mother unpacked the picnic basket while I gathered wood.</td>
<td>C. Compound sentences</td>
</tr>
</tbody>
</table>

2). \textit{The supply types items}, which include completion items and essays items.\textsuperscript{37}

\begin{itemize}
  \item [a)] \textbf{Completion}
\end{itemize}

The completion is a supply-type test item, which can answer by a word, phrase, number or symbol. The completion item is presented in the form of an incomplete statement.

Completion test is a special form of short answer item, it may be defined as a sentence in which certain important words or phrases have been omitted or it directs him to respond to a question by writing his answer in the blank space provided, where only one single answer will be correct. They are likely to measure rote memory rather than understanding.

\textsuperscript{35} Wilmar Tinambunan, \textit{Op. Cit.}, p. 65
\textsuperscript{36} J. Stanley Ahmann & Marvin D. Glock, \textit{Op. Cit.}, p.97
Another name for completion test is fill-in type of question. The differences are the type of question in fill-in is a story and completion is not. Completion test can be used effectively to measure the recall of terms, dates, names and generalization.

According to Gronlund the construction of good short answer or completion item as follow:

State the item so that only a single, brief answer is possible.

1. Start with a direct question and switch to an incomplete statement only when greater conciseness is possible.
2. The words to be supplied should related to the main point of the statement.
3. Place the blanks at the end of the statement.
4. Avoid extraneous clues to answer. The use of the indefinite article ‘a’ or ‘an’ at the end of an incomplete statement is apt to provide a clue to answer.
5. For numerical answer, indicate the degree of precision expected and the units in which they are to be expressed.

According Robert L. Thorndike and Elizabeth Hagen, there are many things concerning completion items, they are:

1. Beware of indefinite or “open” completion items.
   Example:
   The man whose case won freedom of the press for our country was …(Zenger)…

2. Omit only key words.
   Example:
   The constitutional Conventional …(met)… in Philadelphia in 1787.

3. Don’t leave too many blanks in statement.
   Example:

---

The …(ordinance)…of …(1787)… provided for the …(admission)…of …(new states)…

(4). Blanks are better put near the end of statement rather than at the beginning.

Example:
A …(tariff)…is a tax on goods imported into a country.

(5). If the problem requires a numerical answer, indicated the units in which it is to be expressed.40

b) Essay Test

The essay test is usually called as a subjective test because we will find the subjectivity of the teacher in scoring. This subjectivity is caused by the variation of the students’ answer. So, the level of correctness and the error of students’ answer also varieties.

According to Wilmar tinambunan, the essay test is the most complex of supply type items. It demands that the students compose a response, often extensive to a question for which no single response or pattern of response can be cited as correct to the exclusion of are the answer.41

---

4. The Characteristic of a Good Test

The test can be considered as a good test, if it has the certain qualification or the certain characteristics. The characteristics of a good test are:

a. Validity

Validity is the important characteristics of a good test. There are four types of validity, they are: 42

1). Content validity is evaluated by showing how well the content of the test samples the class of situations or subject matter about which conclusion are to be drawn. It is especially important in the case of achievement and proficiency measures.

2). Face validity is not really a type of validity at all, but it appears superficially to measure. A test considered face validity because the test should not “be right”; it is also “look right”.

3). Criterion-related validity, it can be divided into two parts namely: 1). Concurrent validity where the test scores correspond to already accepted measures of performance or status made at the same time, 2). Predictive validity showing how well predictions made from the test are confirmed by evidence gathered at some subsequent time.

4). Construct validity is evaluated by investigating what psychological qualities a test measures, or in other words, by demonstrating that certain explanatory constructs account for performance on the test.

b. Reliability

Drs. Wilmar Tinambunan stated that: “Reliability refers to the consistency of test scores. That is, how consistent test scores or other evaluation results are

---

from one measurement to another”. In the other hand, the writer can say that the
test is reliable if the test can direct the suitable results.

According to Anne Anastasi: “Reliability refers to the consistency of
scores obtained by the same persons when reexamined the same test on different
occasion, or with different sets of equivalent items or under variable examining
condition.”

c. Objectivity

Objectivity of the test is determined by level of equal scores, which is
gotten from the test, although evaluate for values that test (there is no personal
aspects that is influence the evaluation).

d. Usability or Practicality

Usability may be available in more than one form at more than one grade
level. It may be easy to administer properly and to score, or have scored, quickly.
Its cost for student tested may be lower than most other similar tests. Meanwhile,
Robert L. Thorndike and H. P. Hagen said, “Practicality is concerned with a wide
range of factors of economy, convenience and interpretability”.

The writer will not explain all characteristics of good test, therefore only
about validity because it should be conformed by the title of this thesis.

C. Validity

1. Definition of Validity

The test can be a good test if it has certain qualification or certain
characteristics, they are validity, reliability, objectivity and practicability.

\[\text{\textsuperscript{43}Wilmar Tinambunan, Op. Cit., p. 14}\]
\[\text{\textsuperscript{45}Robert L. Thorndike & Elizabeth Hagen, Op. Cit., p. 56}\]
Validity is the most important characteristics of a good test. J Stanley Ahmann and Marvin D. Glock say, “Validity is clearly the most important characteristics of measuring instrument.”46

To know more about validity there are many experts describe about it. They are:

a. Validity is the extent to which an instrument measures that which it supports to measures.47

b. Validity is the degree to which provides us with simple of he behavior that we wish to measure.48

c. Validity is the degree to which it measures what is intended to measure.49

d. Validity is the extent to which a test measures the ability or knowledge that it is purported to measures.50

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

There are several types of validity and some writers use different labels. However, they are similar in procedures and principles. According to Marry Finnachiaro and Sidney Sako, validity is divided into Content validity, Concurrent validity, Predictive validity and Construct validity.51 H. H Remmers and L. L. Gage stated that validity is divided into content validity, concurrent validity, predictive validity and construct validity.52 The other label come from J. Stanley Ahmann and Marvin Glock who stated that validity is divided into: content validity, criterion related validity and construct validity.

47Sawrey and Telford, Educational Psychology, (Boston: Allyn and Bacon, Inc) p. 392
Based on the statement problem so, the writer will explain about content validity only.

2. **Content Validity**

The relevant type of validity in the measurement of academic achievement is content validity. Content validity is validity that talks about the content of the test or a portion of the course represents the content of the entire course.

According to Norman E. Gronlund content validation is the process of determining the extent to which a set of tasks provides a relevant and representative sample of domain of tasks under consideration.53

Kenneth Hopkins defined that content validity of a test is the degree to which the items of that test are a representative sample of the content universe and/ or behavior of the domain being assessed.54

J. Charles Anderson, Caroline Clapham and Dianne Wall said that content validity is the representativeness or sampling adequacy of the content-the substance, the matter and the topics-of measuring instrument.55

Other statement comes from Marlen H. Henerson “Content validity refers to representativeness of the sample the constructs various areas that give each its proportional emphasis.56 Basically, then, content validity is concerned with how well the sample of the test tasks represent the domain of tasks to be measured.

Content validity is also called curricular validity, because the materials that will be tested are based on curriculum. So the way to know content validity is by looking the objective of the curriculum and the material of the test. Dr. Suharsimi Arikunto said that: sebuah tes dikatakan memiliki validitas isi apabila

---

mengukur tujuan khusus tertentu yang sejarah dengan materi atau isi pelajaran yang telah diberikan dan materi itu tertera dalam kurikulum (A test can be said has content validity if it measures the objective instructional according to the materials have been given and it included in curriculum)57 according to Chabib M. Thoha, MA: “Cara untuk menguji validitas isi adalah dengan pendekatan rasional, yaitu membandingkan antara kisi-kisi soal dengan butir soalnya (The way to test the content validity is by rational approach that is comparing the table of specification with the test item).58 While, Norman E. Gronlund said that content validity could be attempted by comparing the test tasks to the test specifications describing the tasks domain under consideration.59

a. Some Factors That Tend to the Influence the Validity of the Results:

1) Unclear direction.
2) Reading vocabulary and sentence structure too difficult.
3) Inappropriate level of difficulty of the test items.
4) Poorly constructed test items.
5) Ambiguity.
6) Test items inappropriate for outcomes being measured.
7) Improper arrangement of items.
8) Identifiable pattern of answers.60

59Norman E. Gronlund, Op. Cit, p. 68
60Norman E. Gronlund, Op. Cit, p. 87-88
b. The Technique of Determine the Content Validity

In Essential of Psychological Testing, the procedure of determining the content validity is comparing the items logically to the content supposed to be measured. Meanwhile, according to Drs. M. Chabib Thoha, MA: “Cara untuk menguji validitas isi adalah dengan pendekatan rasional, yaitu membandingkan antara kisi-kisi soal dengan butir soalnya” (The way to test content validity is by rational approach that is by comparing the table specification with the test item).  

Therefore, the way to know content validity is by comparing the curriculum, table specification and the tests.

---

CHAPTER III
METHODOLOGY AND RESEARCH FINDINGS

A. Methodology of Research

1. Place and Time

The writer did an observation on Tuesday, May 13, 2009. Then, the writer asked the English summative test paper of first year of students of Madrasah Tsanawiyah Pembangunan UIN.

2. The Population and Sample

In this research, the writer took the population from the first students of Madrasah Tsanawiyah Pembangunan UIN, Jakarta as a sample. The whole population of the first class is about 252 students, which is divided into six classes.

Because the population is homogenous, the sample is taken only one class. There are 42 students in 1a classes. The writer used a purposive random sampling to get representative data.

3. Analysis Data

The data in this research are analyzed by making comparison between the English summative tests with the English syllabus (the learning objectives, topic and subtopic, and functional skills) and calculating the percentage of each comparison using the formula:¹

¹ Anas Sudjiono, Statistik Pendidikan, (Jakarta:PT.Raja Grafindo Persada, 2005), p.43
\[
P_c = \frac{f}{N} \times 100\% \]

Where,

- \(P_c\) = Percentage of Content Validity
- \(f\) = Total of the Test
- \(N\) = Total of the Material should be tested

The classification is:

- Conformity if 76% to 100%
- Enough if 56% to 75%
- Low if 40% to 55%
- No Conform if < 40%

4. Instrument of Research

The instrument of this research is syllabus, which contains a number of proportions derived from the English curriculum 2006. The writer compares the proportion of the task and what required in the English syllabus by putting the items of the research variables in the specific columns.

5. The Unit of Analysis

The unit of analysis is the English summative test of semester one for the first year students of Madrasah Tsanawiyah Pembangunan UIN Junior High School Jakarta and its syllabus. This syllabus develops the principles of English curriculum 2006.

---

6. The Procedure of Research
   a. Preparing the objects that will be analyzed.
   b. Designing the table specification (the learning objectives, topic and subtopic, and functional skills) accordance to the English syllabus.
   c. Making classification of conformity.
   d. Analyzing the data by compared the proportion of English summative test and table specification.
   e. Making identification of the areas part form the most valid by make percentage. The areas start from the most valid with the high percentage.

B. Research Finding
   1. The Description of Data

      The type of the test that is studied by the writer is the summative test. The summative test is UAS (Final Semester Examination) semester 1 for the first year students of Madrasah Tsanawiyah Pembangunan UIN, Jakarta, for the academic year 2008-2009. It was prepared and made by the English teacher of Madrasah Tsanawiyah Pembangunan UIN, Jakarta. Meanwhile, the kinds of test items are multiple choice and completion items. The total number of test items is 50 (45 test items are multiple choice, 4 items are completion and 1 test items is jumbled words). Each item consists of the stem and 4 options, which include the key answer and the others, are the distracters. The test was held on Monday, 15 December 2008 with the total time that is given to the testee for answering the whole items within 90 minutes.

      In more detail the writer compares the summative test and test specification with the English syllabus by using table, (see table 3.1, 3.2, 3.3).
Table 3.1: Table of conformity between the summative test and the learning objective of the English Syllabus

<table>
<thead>
<tr>
<th>The learning objective of English syllabus</th>
<th>Test Specification</th>
<th>In the test no</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Reading</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Find the ideas of reading text.</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>b. Find the explicit topic of reading text.</td>
<td>√</td>
<td>34</td>
</tr>
<tr>
<td>c. Find the implicit topic of reading text.</td>
<td>√</td>
<td>10, 11, 36</td>
</tr>
<tr>
<td>d. Find the detail information.</td>
<td>√</td>
<td>8, 37</td>
</tr>
<tr>
<td>e. Find the explicit supporting details.</td>
<td>√</td>
<td>2, 3, 6, 35</td>
</tr>
<tr>
<td>f. Find the implicit supporting details.</td>
<td>√</td>
<td>1, 7</td>
</tr>
<tr>
<td>g. Interpret the meaning of word, phrase and sentence based on the context.</td>
<td>√</td>
<td>9, 12, 23, 26, 33, 44</td>
</tr>
<tr>
<td>h. Get the happiness.</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2. Writing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Change the simple sentence with the correct spelling a correct punctuation.</td>
<td>√</td>
<td>29, 47, 50</td>
</tr>
<tr>
<td>b. Write the simple sentences which dictated by teacher.</td>
<td>√</td>
<td>46, 48</td>
</tr>
<tr>
<td>c. Complete the simple dialogue or short paragraph (narration or description)</td>
<td>√</td>
<td>4, 5, 13, 18,19,20 21, 42, 49</td>
</tr>
</tbody>
</table>

**Total = 11**

**Total = 9**

The writer focused on the learning objectives of English syllabus. It is divided into two aspects, reading and writing. For reading aspect, the syllabus explores the learners ability to master the reading comprehension based on the
passage. In writing aspect, the students should be able to change, write and complete the simple sentence or simple dialogue.

The writer found that there are 11 items learning objectives to be tested, but only 10 items learning objectives are tested, so:

\[
Pc = \frac{f}{N} \times 100\% \]

Where,

- \(Pc\) = Percentage the comparison of English summative test with GBPP
- \(f\) = Total of the items which appear (which is looked for the percentage)
- \(N\) = Total of the items which have been determined

\[
Pc = \frac{9}{11} \times 100\% = 81.8\% \]

Table 3.2:
Table of conformity between the summative test and the topics and subtopics of the English Syllabus

<table>
<thead>
<tr>
<th>Topics</th>
<th>Sub topics</th>
<th>Test Specification</th>
<th>In the test no</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self identity</td>
<td>A. Introduction</td>
<td>√</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>B. Personal Appearance</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>School life</td>
<td>A. At the School</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Topic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>-----</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>B. In the classroom</td>
<td>√</td>
<td>12, 22, 23, 42, 43</td>
<td></td>
</tr>
<tr>
<td>C. in the library</td>
<td>√</td>
<td>35, 36, 37</td>
<td></td>
</tr>
<tr>
<td>D. In the canteen</td>
<td>√</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>E. In the laboratory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F. School garden</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G. School activity</td>
<td>√</td>
<td>1, 2, 3</td>
<td></td>
</tr>
<tr>
<td>H. School rules</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I. Scout</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family life</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Family members</td>
<td>√</td>
<td>10, 21, 49</td>
<td></td>
</tr>
<tr>
<td>B. House and tools</td>
<td>√</td>
<td>31, 32, 40</td>
<td></td>
</tr>
<tr>
<td>C. Family activities</td>
<td>√</td>
<td>30, 46</td>
<td></td>
</tr>
<tr>
<td>D. Daily activities</td>
<td>√</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>E. Neighborhoods</td>
<td>√</td>
<td>8, 33, 34, 48</td>
<td></td>
</tr>
<tr>
<td>F. Pet</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G. Wealth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Profession</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. In the work place</td>
<td>√</td>
<td>11, 14, 19</td>
<td></td>
</tr>
<tr>
<td>B. Kinds of work</td>
<td>√</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>C. Situation of work</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Total = 21**  **Total = 12**

* Note : item no 39 do not include in the first semester.

The table 2 shows that all topics and subtopics of English syllabus semester one. There are 4 topics like Self identity, School life, Family life and Profession. Self identity consists of 2 subtopics. The school life contains 9 subtopics. For family life collects 7 subtopics. And the last, profession consists of 3 subtopics altogether.
The writer found that there are 4 topics with the 21 sub topics must be tested but only 4 topics with the 13 are tested, so:

\[
Pc = \frac{f}{N} \times 100\%
\]

Where,

- \(Pc\) = Percentage the comparison of English summative test with GBPP
- \(f\) = Total of the items which appear (which is looked for the percentage)
- \(N\) = Total of the items which have been determined

\[
Pc = \frac{12}{21} \times 100\% = 57.1\%
\]

### Table 3.3:

Table of conformity between the summative test and the English functional skill of the English Syllabus

<table>
<thead>
<tr>
<th>No</th>
<th>English functional skill of the English syllabus</th>
<th>Test Specification</th>
<th>In the test no</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Comprehend and Express</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>Self introduction</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2.</td>
<td>Introduction someone to the others</td>
<td>√</td>
<td>47</td>
</tr>
<tr>
<td>3.</td>
<td>Greeting in the first meeting/ goodbye</td>
<td>√</td>
<td>4</td>
</tr>
<tr>
<td>4.</td>
<td>Thanks</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Based on the table above, the syllabus explores the learner ability to comprehend and express the English functional skill of the syllabus. There are so many items not included like self introduction, thanks, like/ dislike, imperative,
choice, existence someone / something, explanation about someone, past event and obligation.

The writer found that there are 23 items of functional skill must be tested but only 14 items of functional skill are in the test, so:

\[
Pc = \frac{f}{N} \times 100\%
\]

Where,

- \(Pc\) = Percentage the comparison of English summative test with 1994 GBPP
- \(f\) = Total of the items which appear (which is looked for the percentage)
- \(N\) = Total of the items which have been determined

\[
Pc = \frac{14}{23} \times 100\% = 60.9\%
\]

2. The Interpretation of Data

Before interpreting the data, the first step which be done by the writer is making the table result as follow:

<table>
<thead>
<tr>
<th>TOPIC</th>
<th>SUBTOPIC</th>
<th>RESULT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content Validity</td>
<td>The learning objective</td>
<td>81.8 %</td>
</tr>
<tr>
<td></td>
<td>The topic and subtopic</td>
<td>57.1 %</td>
</tr>
<tr>
<td></td>
<td>The functional skill</td>
<td>60.9 %</td>
</tr>
</tbody>
</table>
The categories of the formula are as follow:

Conformity if 76 % to 100%
Enough if 56 % to 75%
Low if 40 % to 55%
No Conform if < 40%

Based on the table above, it can be said that the test has enough content validity because the result is 66.6 % (Enough). The way is:

\[ Pc = \frac{81.8 \% + 57.1 \% + 60.9 \%}{3} = 66.6 \% \]

The writer can interpret that this summative test has enough content validity that can be seen from the description above and the result is 66.6 % of the content validity.
CHAPTER IV
CONCLUSION AND SUGGESTION

A. CONCLUSION

Before taking a conclusion, the writer passes the process research into two stages. The stages are studying the theoretical principles of the content validity and comparing the content of the object of study – the English summative test with the curriculum used by Madrasah Tsanawiyah Pembangunan UIN, Jakarta. The result of the research is 66.6%. The writer concludes that the test has enough content validity. It means that the test is valid because almost all the test items (content of the test) of the semester one conform to the 2006 English curriculum. The sample test items represent the domain of content to be measured, although there is an item which should be given at the second semester, like no 39.

B. SUGGESTION

Before making test the test maker be sure to make a good preparation in order to get the best result in writing the test. The writer would like to suggest the test makers to:

1. Take the curriculum or syllabus into consideration in planning the test.
2. Take the proper sample of the materials as suggested by the curriculum.
3. Examine the items to get more valid and reliable test by trying the items.
ACKNOWLEDGEMENT

In the name of Allah, The Beneficent and Merciful. All praise to be Allah, the Lord of the universe for the health and strength that enables the writer in finishing her skripsi. Peace and blessing be upon our prophet Muhammad SAW, his families, his relatives and his followers.

In this occasion, the writer would like to express her great honor and acknowledgement to Mr. Farkhan as the advisor to guide her in writing her skripsi and beloved Mom (T I T I H), Father (Makmun. Az), Sisters (Eny, Indah and Yuni) who always gave support motivation and encouragement to fulfill the study.

Her gratitude also goes to:

1. Prof. Dede Rosyada, M. A, The Dean of Tarbiyah and Teacher’s Training Education.
2. Drs. Syauki, M. Pd. & Neneng Sunengsih, S. Pd, The head and secretary of Tarbiyah Faculty.
3. Djamaludin, S. Pd, The Headmaster of Madrasah Tsanawiyah Pembangunan UIN Jakarta, for his permission for the writer to do this research.
5. Lectures of The Study of English Education, The Faculty of Tarbiyah and Teacher’s Training, Islamic State University of “Syarif Hidayatullah”, Jakarta.
6. Sofwan Tarmizi, for his support that never stop. Loving and carrying give her the full power to complete the study.
7. The writer’s best friend (Suci Ningrum, Megasari Nurul I, Erna Satiyah and d’waffle band) for their help in finishing the research maximally without feeling exhausted.

May Allah the almighty bless them all for their generosity and kindness, Amen.

The Writer
# TABLE OF CONTENT

ACKNOWLEDGEMENT ............................................................ iv

TABLE OF CONTENT .......................................................... vi

CHAPTER ONE INTRODUCTION ............................................. 1
   A. Background of the Study .......................................... 1
   B. Statement of the Problem ....................................... 4
   C. Scope and Limitation of the Problem ......................... 4
   D. Purpose and Use of the study .................................. 4
   E. Method of Research ............................................... 5

CHAPTER TWO THEORETICAL FRAMEWORK ......................... 6
   A. Syllabus ................................................................. 6
      1. Definition of Syllabus.......................................... 6
      2. Purpose of Syllabus .......................................... 7
      3. Component of Syllabus ...................................... 8
      4. Format of Syllabus .......................................... 8
   B. Test ................................................................. 12
      1. Definition of the Test ...................................... 12
      2. Purpose of the Test ....................................... 13
      3. Types of Test and Test Items ............................. 14
         a. The Types of Test .................................... 14
         b. The Types of Test Items ............................. 16
      4. The Characteristic of Good Test ......................... 24
   C. Validity ............................................................... 25
      1. The Definition of Validity ................................ 25
      2. Content Validity ............................................. 27
         a. Some Factors That Tend to the Influence
            The Validity of the Results .......................... 28
         b. The Technique of Determine the Content
            Validity.................................................... 29
REFERENCES


Arikunto, Suharsimi, Drs, 2008, *Dasar-dasar Evaluasi Pendidikan Edisi Revisi* Jakarta: Bumi aksara


Purwanto, M. Ngalim, Drs, *Prinsip-Prinsip dan Teknik Evaluasi Pendidikan*, Bandung: PT Remaja Rosda Karya


Sawrey, & Telford, *Educational Psychology*, Boston: Allyn and Bacon, Inc


