AN ANALYSIS ON CONTENT VALIDITY OF ENGLISH SUMMATIVE TEST
(A Case Study of Second Grade at SMA Negeri 87 Jakarta, South Jakarta)

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)

By:
MOH. CAHYONO ADHI NUR
NIM. 203014001575

Approved by:
Advisor
Dr. H. M. Farkhan M.Pd
NIP. 150 299 480

ENGLISH EDUCATION DEPARTMENT
FACULTY OF TARBIYA AND TEACHERS’ TRAINING
STATE ISLAMIC UNIVERSITY SYARIF HIDAYATULLAH
JAKARTA
1430 H / 2010 M
ABSTRACT

Evaluation plays an important role in some activities, especially in terms of education. Evaluation gives information about how successful an effort has been. The information gain through this evaluation will be very useful to make improvement in the future. One of the instruments of the evaluation in education field is a test.

A test is supposed to be well arranged so that it can be used effectively. To be said a good test; it has to fulfill the characteristic of the good test, these are validity, reliability and practically. It is valid if the test can measure what is supposed to measure. It can be reliable if the result from the test is the same even thought the test is administrated to the same students for several times. A test can be practical if it easy to do and administer.

The items analysis can give some feedback for the students and teacher as the test maker, it can determine whether an functions as teacher intends, feedback to students about their performance and as basis for class discussion, feedback to the teacher about pupil difficulties, area for curriculum improvement, revising the item and improving item writing skills.

The writer held the research at second grade student of SMA Negeri 87 Jakarta which is located on Jl. Mawar II Kel. Bintaro Pesanggrahan Jakarta Selatan, phone number (021) 73881969.

Key words: evaluation, test, item analysis.
TABLE OF CONTENTS

| ABSTRACT | ........................................................................................................ i |
| ACKNOWLEDGEMENT | ................................................................................................... ii |
| TABLE OF CONTENT | ................................................................................................... iv |

CHAPTER I  INTRODUCTION

A. The Background of Study ............................................. 1
B. The Research Question ................................................. 5
C. The Limitation of Problem ............................................ 6
D. The Objective of Study ................................................. 6
E. The Significance of Study ............................................. 6

CHAPTER II  THEORETICAL FRAMEWORK

A. Evaluation
   1. The Meaning of Evaluation.................................... 7
   2. The Function of Evaluation................................... 8
   3. The Type of Evaluation........................................ 9
B. Test .............................................................................. 10
   1. The Definition of Test................................. 10
   2. The Purpose of Test .................................... 11
   3. The Type of Achievement Test .................... 12
      a. Formative Test ........................................ 12
      b. Mid Test .............................................. 13
      c. Summative Test ..................................... 13
   4. Types of Test Item........................................... 14
      a. Subjective Test..................................... 14
         1. Essay............................................ 15
         2. Short Answer Question..................... 16
         3. Completion ................................. 16
CHAPTER III THE IMPLEMENTATION OF THE RESEARCH

A. RESEARCH METHODOLOGY

1. Method of The Study ................................. 28
2. The Object of the Research ........................... 29
3. The Time and Location of Research ............. 29
4. Technique of Data Collecting ....................... 29
5. Technique of Data Analysis ........................ 29

B. RESEARCH FINDINGS

1. Description of Data .................................... 30
2. The Analysis of The Data ............................ 31
   a. The Functional Skills ............................... 31
3. Data Interpretation ................................. 38

CHAPTER IV CONCLUSION AND SUGGESTION

A. Conclusion ........................................... 39
B. Suggestions .......................................... 40
APPENDIX

Appendix 1 : Items test
Appendix 2 : Syllabus
Appendix 3 : Surat Keterangan Penelitian
Appendix 4 : Surat Bimbingan Skripsi
Appendix 5 : Surat Permohonan Judul Skripsi
LIST OF TABLE

1. Table 2.1 : Three types of validity 21
2. Table 3.1 : List of Functional Skills 30
3. Table 3.2 : The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus 31
4. Table 3.3 : The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus 32
5. Table 3.4 : The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus 32
6. Table 3.5 : The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus 33
7. Table 3.6 : The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus 34
8. Table 3.7 : The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus 34
9. Table 3.8 : The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus 35
10. Table 3.9 : The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus 35
11. Table 3.10 : The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus 37
CHAPTER I
INTRODUCTION

This chapter presents and discusses the background of the study, the research question, the limitation of the study, the objective of the study and the significance of the study.

A. The Background of The Study

Nowadays, English is the first foreign language in our country, and English language is studied by all students not only in primary school but also at the university. The students realize that English is very important, because by using English language, they can communicate to the others communities, they can get information and new knowledge. English is an international language. It functions as communication instrument in global era and as an instrument of the students in all sort fields in order to be intelligent citizens, skillful, and able to face the competition era.

The objective of teaching English in the Sekolah Menengah Atas/Madrasah Aliyah are to develop the communication skills especially in oral and written skills (listening, speaking, reading and writing), to grow the students’ interest and the most important of English language as the one of foreign language as a main tool to study, to develop student’s understanding the relationship
between language and culture. So, the students understand their culture and take part on it.¹

_Pusat Kurikulum_ also states that the purpose of English teaching at the second grade of _Sekolah Menengah Atas/Madrasah Aliyah_ its to encourage the student to communicate in oral or written and can use the language fluently and accurately in interactional or monologue concept which are concern in descriptive, narrative, recount, and procedure text with kind of interpersonal ideational and simple textual expressions.²

When the students study a language, they have to know two skills, they are receptive skills (listening and reading) and productive skills (speaking and writing).

To measure their students’ skills in the teaching-learning process, the teachers need to hold an evaluation. It is one of the main aspects in teaching-learning activities process. It is intended to know the success and effectiveness (strength and weakness) of teaching-learning activities. It is integral part of the instructional program.

Through the evaluation, teachers are able to find failure of a method and the curriculum. It also measures and diagnoses students’ ability and their achievement in mastering the lessons effectively.

According to Prof. Drs. Sudjiono in his book _“Pengantar Evaluasi Pendidikan”_ that evaluation can be meant as everything which has correlation with education activities.³ Norman E. Gronlund explains that evaluation is a systematic process of determining the extent to which pupils achieve learning objective.⁴

---

¹Pusat Kurikulum-Dit PLP,Wisata, Jakarta, 2003, p. 7-8  
²Sarkim, T. Drs., _KTSP Lahir Prematur, Perlu Penegakan Otonomi Sekolah_, Educare Nomor 4 IV Juli 2007 p. 7-8  
Based on the statements above, evaluation has to arrange systematically and relatedly to the curriculum so that we can get an obvious and comprehensive analysis of the students’ achievement.

By having an evaluation, the improvement in the next lesson planning can be upgraded. So, teaching-learning process will run effectively without any overlapping.

Making a good test needs careful arrangements. Test as an instrument has to have a good quality, because the quality of the test will influence the result of the test. If the test is good, the result will provide the right information to be used by the teacher in making accurate decision to the students’ achievement.

Prof. Dr. Anas Sudjiono states that an instrument can be valid if it can measure a represent the data of variable what their observed accurately. The high or low the data of instrument can show us how far the data which are collected are no to draw a side from the validity concept.”\(^5\) So, validity is important to make a good test, and then it must reflect some subject matters, knowledge, and skills.

The good test is also have the reliability, the test is said to be reliable if the scores do not change much when the test is re-administrate.\(^6\) But in reality, when the teachers do an evaluation, they find some problems and one of them is some of their students get low scores. Based on the reality, we can assume that was happened because the items test didn’t contain one of the characteristics of a good test, or the teacher’s items test didn’t in line with their curriculum or may be the teachers didn’t deliver the material to their students completely.

In the same situation, the writer also found the same problem in th English summative test at the second grade of SMA Negeri 87 Jakarta. He found some the students of that school got low scores in their English Summative test. He knew from their answer sheets. The writer tried to investigate this problem. The result was some of them felt the test was too easy, so they got high score. For the others

students, they felt the test were too difficult, they couldn’t do the test completely and they didn’t understand the content of the test. So, they got low scores.

The writer guessed because the items test were invalid, maybe the items were unsuitable to the one of the characteristics of a good test, or maybe their students’ didn’t study hard to do English Summative Test. So, the writer is interested to investigate the problem.

In this case, the writer will observe validity especially content validity. It is important to know that validity is not measured directly, but is inferred from the data. Thus it is a judgment to whether evidence of validity is adequate marginal or unsatisfactory. In similar manner, it is also appropriate for our information. Remember that validity is a matter of degree. A test is almost never completely valid, nor is it usually entirely invalid.

Content validity is important thing for the teacher. The teacher who gives an examination that covers the materials and objectives of instruction within her classroom has probably given a test that has content validity.

It is important to remember that content validity is especially vital for measurement of adjustment based on observation. A standardized achievement test is judged as having content validity when it’s content the instructional materials.

In the making of the test process, the test can be classified into standardized test and teacher made test. Test authors and publishers who procedure test for national use strive to determine the generally accepted educational aims of instruction in the area in which a test is to be constructed. They do this by surveying textbooks, class syllabi and subject-matter to determine the entire range of skills, knowledge’s, process, and behaviors covered in successive grade levels in subject-matter areas. Thus constructing a nationally standardized is a monumental task. If we do not want test maker dictating curricula, then you should carefully check that specific test sample the skills and
objectives your school wishes to develop in students and to evaluate by test scores.\textsuperscript{7}

The content observed of such as the learning view objectives of reading, writing skills, the functional skills, and themes/sub-themes of the KTSP of that school.

Based on Indonesian schooling system, the students’ achievement test consists of three activities, these are: formative test, mid test and summative test.

As Benjamin Bloom states that the formative test is given during the course of instruction; its purpose is to show which aspects of the chapter the students mastered and where remedial work necessary. The mid test is given at the middle of a school term or at the middle of semester. The summative test is given in the end of semester, this test is to measure the students’ achievement for all chapters and the summative test is carried out after some units of lessons have been running at the end of a term, usually after one semester.\textsuperscript{8}

So based on the explanation above, the writer will investigate more about evaluation especially in content validity of English summative test for the second grade of SMA Negeri 87 Jakarta it is located on Jl. Mawar II Kelurahan Bintaro South Jakarta.

**B. The Research Question**

Based on the description above, the research question of this research is the material tested in English summative test for the second grade of SMA Negeri 87 Jakarta in line with their KTSP?


\textsuperscript{8}Lynn Lyons Morris and Carol Taylor Fitz-Gibbon, *Program evaluation Kit : How to Measure Achievement*, University of California, Los Angeles, 1978, p. 106-107
C. The Limitation of the Study

The writer limits the problems to the conformity level of the content validity on The English summative test with English KTSP curriculum of SMA Negeri 87 in terms of functional expression in speaking skills on 2nd grade consist of 10 functional skills.

D. Objective of the Study

The objective of the study is to describe about the material tested in English summative test for the second grade of SMA Negeri 87 Jakarta in line or not suitable with their English KTSP.

E. The Significance of the Study

The result of this research is to give some suggestions for (a) the English teacher, (b) the school principal, and (c) further researchers. They are described as follows:

a) The teacher who makes or arranges the summative test to provide better insight on how to make the better English summative test.

b) The result of this study also useful for the school principal as a decision maker at school, he also has responsibility to check how well the English summative test designed by the English teacher.

c) Finally, the result of this study can be used as a reference for the next researchers who are interested in developing similar study.

Furthermore, the result of this study can bring the positive impact in evaluation process at second grade students.
CHAPTER II
THEORETICAL FRAMEWORK

This chapter presents and discusses the theoretical framework which consists of: the definition of the evaluation, the definition of the test, the definition of content validity, the Kurikulum Tingkat Satuan Pendidikan (KTSP) of English and the conformity of content validity.

A. Evaluation
1. The Meaning of Evaluation

The term evaluation tends to be used somewhat ambiguously in relation to other terms such as assessment and testing. Evaluation can make assessment instruments (including test), but it is not limited to such form of information gathering. It may include, for example the use of unstructured interviews. Likewise, assessment instruments (including test) can be used for purpose other than evaluation, such as to measure individual language ability in order to test a research hypothesis concerning language acquisition.¹

There are some definitions about the evaluation and the assessment to differentiate both of them.

Evaluation is the systematic process of collecting and analyzing data in order to determine whether, and go to what degree, objective have been, or are being achieved, evaluation is also used to make decision.²

Evaluation is defined as the systematic attempt to gather information in order to make judgment decisions.\(^3\)

Assessment is the process of gathering and discussing information from multiple and diverse sources in order to develop a deep understanding of what students know, understand, and can do with their knowledge as a result of their educational experiences; the process culminates when assessment results are used to improve subsequent learning.

Assessment is the process of collecting, analyzing, and interpreting information about academic performance related to educational goals. Assessment, as opposed to test, typically employs a variety of procedures for evaluating student work and learning.\(^4\)

Based on the definitions above, the writer concludes that the evaluation are some activities in systematic procedures for collecting and analyzing and interpreting information of academic related to the teaching-learning goal.

So based on the statement above, evaluation should be arranged systematically and related to the curriculum, so that we can get an obvious and comprehensive analysis of the students' achievement. By having an evaluation, the improvement in the next lesson planning can be upgraded by focusing on what student really need to acquire so that the teaching-learning process will run effectively without any overlapping.

Evaluation in our daily life is a natural activity. It is something that can be very formal and informal. It is also something that may not always be made explicit but may actually be undertaken unconsciously.

2. The Function of Evaluation

The function of evaluation divided into two parts:

a. In general

The general functions of evaluation are to measure the achievements, to support a plan arrangement, and it is also to repair or recompiled.

\(^3\) Bryan K. Lycnh, *Language Program.....*, p. 2
\(^4\) “Definition of Assessment at Pittsburg State University” accessed on September 13\(^{th}\) 2009 from [http://www.diknas.com](http://www.diknas.com)
b. In specific

The specific functions of evaluation, they are consists of three views:

- Psychology view
  Based on the psychology view, the function of evaluations are to know the students’ capacity (for the students), and to know the result of the students’ effort (for the educators).

- Didactic view
  Based on the didactic view, the function of evaluation are to encourage the students’ achievements (for the students), and for the educators, the evaluation is used to diagnostic, appointment, selective, guidance and instructional function.

- Administrative view
  Seen from the administrative views, the function of evaluations are to give information, the data result and illustration.5

3. Type of Evaluation

Based on L. R. Gay in his book entitle “Educational Evaluation and Measurement”, he divides evaluation into three types, and they are:

a. Student Evaluation

When we think of student evaluation, we probably think of achievement. Achievement is one of many variables on which a student assessed. Other variable include: aptitude, intelligence, personality, attitudes and interest, to name a few of the major ones. May different kinds of data are collected on each student in order to make a variety of decisions. Test, both standardized and teacher made are administrated, projects, procedures, and oral presentations are rated, and format and informal observation made. The data obtained are recorded on teacher’s records and in guidance files, which include, of course, the ever-popular cumulative folder. Not all of the subsequent decisions relate directly to the student. A teacher uses performance data not only to evaluate student progress but

5Diknas, Pengenalan dan Pengembangan KTSP, p. 9
also to evaluate her or his own instruction. In other words, the process of evaluating students provides feedback to the teacher.

b. School Evaluation

Evaluation of a school involves evaluation of the total educational program of the school and entails the collection of data on all aspects of its functioning. The purpose of school evaluation is to determine the degree to which school objectives are being met and to identify areas of strength and weakness in the total program. The information provides feedback which gives direction to the future activities of the school and results in decisions concerning the allocation of school resources.

c. Curriculum Evaluation

Curriculum evaluation involves the evaluation of any instructional program or instructional materials, and includes evaluation of such factors as instructional strategies, textbook, audio-visual materials, and physical and organizational arrangements. Curriculum evaluation may involve evaluation of a total package or evaluation of one small aspects of a total curriculum, such as a film. Although ongoing programs are subject to evaluation, curriculum evaluation is usually associated with innovation, a new or different approach.6

B. Test

1. The Definition of Test

We have to know that evaluation is an activity, which is done to get the information of learning report and to be used in making educational purpose. According to Jack C. Richard one of methods is a test. As one method for collecting information, it has certain important aspects, they are: content, method, and measurement.7

---

6 Gay, Educational Evaluation.....p. 16-19
The test word is derived from *testum*. It means the plate to keep off the pure metals. Test is a tool or a procedure used to measure and knows something based on the specific clues. If we are talking about test are some terminologies. Firstly is *testee*, are the respondences that are doing the test. Secondly is tester, is someone who is responsible to collect and resume the respondences’ result.\(^8\)

Tests are used for different evaluative needs, and for each need a different method of investigation is necessary to establish validity. Different kinds of tests are used for various measurement purposes. It should be noted, however, that the purposes of many test overlap. Therefore, it is necessary to gather different validity data within one test well as to secure validity information on different type of test.\(^9\)

2. The Purpose of Test

Based on Prof. Drs. AnasSudjiono’s statement on his book “*Pengantar Evaluasi Pendidikan*”, the purpose of the test of evaluations is:

a. To collect the information which can be evidences about the student’s progress after they followed some learning process in a certain time.

b. To measure and identify the effectiveness is using methods in teaching-learning activities in a certain time.

c. To determine the current status of the object the evaluation to compare that status with a set of standards, of criteria and to select an alternative in order to make a decision. There may be only two alternatives (e.g. continue the program or not, adopt the new curriculum or keep the current one)\(^10\)

According to Hughes, there are different purposes that require different kinds of test. Those are:

---


10Sudjiono, *Pengantar Evaluasi*.....p. 24
• To measure language proficiency/regardless of any language courses that candidates may have followed (proficiency test)
• To discover how far the students have achieved the objectives of a course of a study (achievement test)
• To diagnose student’s strength and weakness to identify what they know and what they don’t know (diagnostic test)
• To assist placement of students by identifying the stage or part of teaching programmed most appropriate to their ability (placement test)\(^\text{11}\)

3. Type of Achievement Test
   a. Formative Test

   Formative test is usually given during the learning process; there are some perspectives of the formative test maker: the formative test maker major responsibility concerning achievement is to make progress checks throughout the course of the program and to ensure that students are learning what is expected and keeping to the anticipated pace.

   Another task of the formative test maker may be to conduct short experiment to try out program components or to settle controversies among planners about which version of program implementation will work best. In these situation the formative evaluators selects or constructs test which measure the specific achievement objectives of the components under serenity.

   Because of less rigid data collection requirements, the formative test maker can be more flexible than summative test maker in choosing achievement instrument. The end-of-unit test which often accompany curriculum materials, for instance, are usually suspect as measures of program achievement for the purpose of summative evaluation because such test may too narrowly focus on the particular program materials used. For formative evaluation, however, unit tests are ideal data collection methods, they tell the staff whether students are at least learning what the program materials contain.

---

In general, the formative test maker can assess achievement using whatever method and him immediate audience perceive as providing as providing credible evidence. Of course, if the formative test maker produces measurement instruments to settle controversies among the staff about alternative ways to implement the program, the need for technically defensible measures will again become critical.

The formative test maker may want to collect data what will stimulate discussion among staff members and planners. Monitoring student progress in achieving program objectives can prod the staff to set more reasonable expectation about what the program can accomplish in the time during which the students are exposed to it. By measuring achievement of skills which are not mentioned among the program’s stated objectives, the formative evaluator’s strongest contribution to the program could turn out to be the program did not take into account or the identification of abilities which the program indeed produced by which has not been identified as program objectives.12

b. Mid Test

Mid test is usually given at the middle of learning process. The teachers usually give this test after they finished a half of all materials or chapters in the semester. This test is used to measure the student’s achievement and also their comprehension about the materials during the middle semester of learning process (usually every three months).

c. Summative Test

Summative test is usually given at the end of learning process or school year. There are some perspectives of the formative test maker; the summative evaluator’s primary interest is to find and use instrument which measure whether the program obtained its overall goals. For this reason, the summative test maker must pay close attention to the programs announced and apparent achievement objectives. He interested in designing or purchasing measures which are most

12 Lynn Lyons Morris, *Program Evaluation*...., p. 11-12
sensitive to determining the accomplishment of these objectives. The summative test maker interest in describing the programs overall impact on achievement should extend, as well, to a concern for measuring cognitive gains which seem to be occurring but which program designers have not mentioned among their objectives.

The summative test maker may have additional reasons for measuring achievement. If he/she using an evaluation design to determine the extent to which good outcomes can be attributed solely to the program, then he/she may wish to use an achievement test for assigning students, classroom school to program and comparison groups, this common use of achievement information, called blocking or stratified sampling, assures that the groups produced will be, at the outset, as like in prior achievement as possible.

The summative test maker may receive a request to include in the final report not only a statement of the program’s achievement of its own objectives. But also an indication of how participants in the program now compare in their achievement with people in other school or districts, in these cases, the evaluator needs to select, administer, and report the result of achievement test that provide such normative data.

Because the summative test maker produces a report which could effect important decision about the program’s future, he/she needs to ensure high credibility in the instruments he/she produces of purchases. That is, she must use instruments that he/her evaluation audiences regard as appropriate and accurate.13

### 4. Types of Test Item

Based on the manner of scoring, the type of test item is divided into two general types.

a. Subjective Test

Subjective test is a test where in its scoring requires judgment and evaluation of the scorer. As stated by Arthur Hughes, “If no judgment is

---

13 Lynn Lyons Morris, *Program Evaluation...*, p. 11-12
required on the part of the scorer, then, the scoring is objective…if judgment is called for, the scoring is said to be objective”.14

In subjective test, the students are given a freedom to answer the question in their own words. The subjective tests that are common used in classroom are essay, short answer, and completion.

1. Essay

The essay item is the most complex of supply type item. 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 all the answer.15

The essay test usually consist of questions beginning with or including such directions as discuss, explain, outline, evaluate, define, compare, contrast, and describe.16

Example: Explain the definition of item analysis!

Thus, the distinctive feature of essay question is freedom of response it provides. In answering the question, the students are given freedom to select, relate, and present ideas in their own words. Because of the feature, the essay test usually scored differently by the same person on different occasion.

Meanwhile, there are some advantages of essay items, namely:

1) The essay question is useful for measuring the pupil ability to organize, integrate, and express their ideas.

2) Constructing essay items are relative easy.

3) The possibility of guessing requires less time than other measurement tool such as true false or on multiple choice items.

Besides the above advantages, the essay items have few disadvantages, namely:

1) Scoring essay items is difficult
2) Unreliability of scoring
3) Judgment of students; respond requires much time.
4) Limited sample of total instruction content.

2. Short Answer Question

The short answer item consists of a question, which can be answered with a word or short phrase. For example: An incorrect response included in multiple choice items is called…

Generally, teachers prefer to use the short answer type of question probably because they think it has some advantages. It is relatively easy to construct. It also gives the teacher some opportunity to see how well students can express their thought and it is also not difficult to score or mark than the essay questions. However, it is difficult to phrase the short answer question so that only one answer is correct. This type of question will be more useful only in testing knowledge of facts and quite specific information. Thus, when teachers are going to know the broader description about something, they are better to use the essay form.

3. Completion

The completion item is a written statement that requires the examinee to supply the correct word or short phrase in response to an incomplete sentence, a question, or a word association. Completion test can be used effectively to measure the recall of terms, dates, and names.

This type of test can be used at almost all levels. But it is extremely difficult to phrase the question or incomplete statement so that only one answer is correct, and in making the questions, it may not too many clues are given to answer the questions. Cause if too many clues are given, the items will be too easy, and if an insufficient number of clues are presented,

18 Victor H. Noll, Introduction to Educational..., p. 139.
19 Wilmar Tinambunan, Evaluation of Student..., p. 61.
the item will be ambiguous and may yield several possibility of correct answer.20

b. Objective Test

An objective test item is any item that there is only a single predictable correct answer. The test is scored in such a manner that subjective judgment is eliminated when determining the correctness of pupil’s answer.21 Therefore, whether one teacher or another scores the item, today or last week, it will yield the same score.

The objective test items commonly used in classroom testing are true false, matching, and multiple choices.

1. True False

True false item is referred to alternative response item, the items asks the students to answer with “true” if it conform to the truth or “false” if it essentially incorrect.22

Thus, the item provides the students with a choice of two alternatives, so the students have a possibility to guess the answer and sometimes it will be the right answer. Because of the random guessing to produce the correct answer, the true false tests are become less reliable that the other types of test. However, these items are appropriate for occasional use, for example after the students choose the two alternatives between right and wrong, correct or incorrect, etc, they asked to explain by writing the sentences justifying their response.23

20 Wilmar Tinambunan, Evaluation of Student..., p. 62.
22 Wilmar Tinambunan, Evaluation of Student..., p. 70.
Another advantage of constructing a true false item is that the students are able to respond to more true false items in a given time period than any other selection type items.\textsuperscript{24}
The most commons uses of true 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.

To measure aspect of understanding, that is the ability to recognize cause and affect relationship. This type of item usually contains two true prepositions in one statement, and the pupil asked to judge whether the relationship between them it true or false.

3) To measure the simple aspect of logic as illustrated in the example.

2. Matching

The matching test item consists of two parallel columns with each word, number, or symbol in one column being matched to a word, sentence, or phrase on the other column.\textsuperscript{25} This type of item is employed widely in situation where relationship of more or less similar ideas, facts, and principles are to be examined or judged.\textsuperscript{26} This kind of test is an effective way to test students’ recognition of the relationships between words, definitions, events, dates, categories, example, and so on.\textsuperscript{27}

Matching items are also useful in measuring students’ ability to make association, interpretations or measure knowledge of a series of fact. Besides that, the matching items can be used for a 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{28}

\begin{itemize}
\item[25] Wilmar Tinambunan, \textit{Evaluation of Student...}, p. 64.
\item[26] Victor H. Noll, \textit{Introduction to Educational...}, p. 64.
\end{itemize}
3. Multiple Choice

A multiple-choice item consists of one or more introductory sentences followed by a list of two or more suggested responses from which the examinee chooses one as the correct answer.\(^{29}\)

The multiple choices item can measure a variety of learning outcomes from simple to complex, and it is adaptable to most types of subject matter content. The learning outcomes in the knowledge area that can be measured by the multiple choice items are:\(^{30}\)

1) Knowledge of terminology

For this purpose, pupils are requested to show their knowledge of a particular term by selecting a word that has the same meaning as the given term or by choosing a definition of the term. Special uses of a term can also be measured by having pupils identify the meaning of the terms when used in context.

2) Knowledge of principles

The multiple-choice items can also be constructed to measure knowledge of principles. In this case, the items appear a bit more difficult, this is because principles are more complex than isolated facts.

3) Knowledge of method and procedure

The multiple choice form is also be able to measure the knowledge of method and procedure, such as knowledge of laboratory procedure, knowledge of methods used in problem solving, computational and performance skill.

Some advantages of using multiple choice items are: the multiple choice items are fast, easy and economical to score, they can be objectively so that they will be fairer and more reliable than subjectively scored tests.


\(^{30}\) Wilmar Tinambunan, *Evaluation of Student...,* p. 76.
Besides those advantages, the multiple choice items also have disadvantages such as: the technique of the test only recognition knowledge, so the students have no or little opportunity to express their own idea of a problem, pupils have much time to guess the answer and it may effect on their scores, it is difficult to write successful items, and cheating may be facilitated.\(^{31}\)

c. Definition of Item Analysis

There are several meanings of what item analysis. According to Anthony J. Nitko, in his book, he states that: “Item analysis refers to the process of collecting, summarizing, and using information about individual test items especially information about pupils’ response to item.\(^{32}\)

Each item of the test needs to function properly otherwise, it can be weakening the test. Fortunately, there are some rather simple statistical ways of checking individual items. The procedure is called item analysis.

Item analysis is an important and necessary step in the preparation of good multiple choice test. Because of this fact, it is suggested that every classroom teacher who uses multiple choice test data should know something of item analysis. How it is and what it means.\(^{33}\)

According to J. Stanley Ahman and Marvin D. Glock, “Item analysis is reexamining each test item to discover its strength and flaws”.\(^{34}\)

The objective of the item analysis is to find which of the test item is good or needs to be corrected. As states by Suharsimi Arikunto, “Analisis soal antara lain bertujuan untuk membantu kita dalam mengidentifikasi butir-butir soal yang jelek, memperoleh informasi yang akan dapat digunakan untuk menyempurnakan...

---

33 Jhon W. Other, *Language Test at School*, London: Longman Group, 1979, p. 245

Based on explanation above, the writer concluded that the item analysis is a process of collecting information about pupil’s responses to the items to see the quality of test items. It is very useful for the teachers and also the students. In analyzing the items, the writer analyzes the difficulty level of the items.

5. The Characteristics of a Good Test

Based on Norman E. Gronlund in his book “Measurement and Evaluation in Teaching”, the most essential of the characteristics of a good test can be classified into three main aspects, they are: validity, reliability, and usability.

a. Validity

Validity refers to the extent to which the result of an evaluation procedure serves the particular uses for which they are intended. If the results are to be used to describe pupil achievement, we should like them to represent the specific achievement we wish to describe, to represent all aspects of the achievement we wish to describe, and to represent nothing. Basically, validity is always concerned with the specific use to be made of the result and with the soundness of our proposed interpretations.

When using the term validity, in relation to testing and evaluation, there are a number of cautions to borne in mind:

1. Validity refers to the result of a test or evaluation instrument for a given group of individuals, not to the instrument itself. We sometimes speak of the validity of a test, for the sake of convenience, but it is more appropriate to speak of the validity of the test result, or more specially, of the validity of the interpretation to be made form the result.

2. Validity is a matter of degree. It does not exist on an all or none basis. Consequently, we should avoid thinking of evaluation 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.

3. Validity is always specific to some particular use. It should never be considered a general quality.

There are three types of validity have been identified and are now commonly used in education, they are: content validity, criterion validity, and construct validity. The 3 kinds of validity will present in the table below.

---

### Table 2.1
Three types of validity

<table>
<thead>
<tr>
<th>Type</th>
<th>Meaning</th>
<th>Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content Validity</td>
<td>How well the sample of test represent the domain of task to be measured</td>
<td>Compare the test task to the test specification describing the task domain under consideration.</td>
</tr>
<tr>
<td>Criterion-related Validity</td>
<td>How well test performance predicts future performance or estimates current performance on some valued measures other than the test itself</td>
<td>Compare test score with another measure of performance obtained at a later date (for prediction) with another measure of performance obtained concurrently (for estimating present status)</td>
</tr>
<tr>
<td>Construct Validity</td>
<td>How test performance can be described psychologically</td>
<td>Experimentally determine what factors influence scores on the test.</td>
</tr>
</tbody>
</table>

---

b. Reliability

Reliability refers to the consistency of evaluation results. If we obtain quite similar scores when the same test administrated to the same group on two different occasion, we can conclude that our result have a high degree of reliability from one occasion to another. Similarly, if different teacher independently rate the same student at the same instrument and obtain similar ratings, we can conclude that the result has a high degree of reliability from one
rater to another. Reliability is intimately related to the type of interpretation to be made. For some uses we may be interested in asking how reliable our evaluation result are over a given period time, and for other, how reliable they are over different samples of the same behavior.

In all instances in which reliability is being determined, however, we are concerned with the consistency of the results rather than with the extent to which they serve the specific use under consideration.

c. Practicality

An evaluation procedure must meet certain practical requirements. It should be economical from the viewpoint of both time and money, it should easily administer and scored, and it should provide results that can be accurately interpreted and applied by the school personnel available. These practical aspects of an evaluation procedure can all be included under the heading of usability. The term usability then refers only to the practicality of the procedure and implies nothing about the other qualities present.36

C. Content Validity

L. R. Gray stated that content validity is the degree to which a test measure an intended contends area. It is of prime importance for achievements test. Content validity requires both item validity and sampling validity. Item validity is concerned with whether the test items represent measurement in the intended content are, and sampling validity is concerned with how well the test samples the total content area.

Test scores cannot accurately reflect student achievement if the test does not measure what the students were supposed to learn. While this may seem obvious, content validity is a problem in many evaluations. Many evaluation studies, for example, are designed to compare the effectiveness of two (or more) different curricula. Effectiveness is often defined in terms of final achievement of two

groups as measured by a test. Sometimes the test used is more content valid for one of the groups than for the other. When this happens, final achievement difference may be least partially attributable to the test used and not just curriculum difference.\textsuperscript{37}

Norman E. Gronlund states that content validity may be defined as the extent to which a test measures a representative sample of the domain of task under consideration. In classroom testing, the domains of achievement tasks are determined by the instruction, and establishing content validity involves (1) clearly defining the domain of task to be measured and (2) constructing, or selecting, representative set of test task.

Content validity is important when we wish to describe how an individual performs on a domain of task that the test is supposed to represent. We might, for example, expect pupils to be able to spell the 200 words on a given word list. Because a 200-word spelling test would take too time consuming, we might select a sample of 20 words to represent the total domain of 200 words spelling words. Now, if a pupil correctly spells 80 percent of these 20 words. Thus, we would like to be able to generalize from the pupil’s performance on the sample of spelling words in the test to the domain spelling words spelling words represent. Such interpretations would be valid to the extent that our twenty-word test constituted a representative sample of the 200 words.\textsuperscript{38}

D. The KTSP of English

There some opinions from the experts and observers of education about KTSP:

According to Drs. Nasarius Sudaryo opinion, KTSP gives autonomy to the teacher in curriculum arrangement. Teacher has a freedom to develop indicators which are suitable with the student’s characteristics and their region.

He added that we can do the KTSP has three aspects, they are: indicators, material, and learning activities. Indicators can be arranged as a specific

\textsuperscript{37} Gay, \textit{Educational Evaluation}..., p. 159
\textsuperscript{38} Gay, \textit{Educational Evaluation}..., p. 159
competence and it also very important to decide all contents and in the learning activities. A good material is a material which are able to develop the student’s thinking skill and it is also in line with the student’s progress and suitable with their life. In this case, learning activities have to orientated to the students (Students Centered Learning), it is a good approach because it can stimulate the student’s high thinking level ability, it can give them a freedom to learn about work together and it also encourage the to socialization with everything around them.

( tiga aspek pokok yang menjadi pilar KTSP adalah: indicator, materi dan kegiatan pembelajaran. Indicator dapat dirumuskan sebagai kompetensi yang paling spesifik. Indicator ini yang akan sangat menetukan keseluruhan isi dan proses pembelajaran. Materi dapat dikategorikan baik jika materi yang dimaksud mampu mengembangkan kemampuan berfikir tingkat tinggi siswa serta selarat dengan irama perumbuhan anak dan dekat dengan kehidupan siswa. Kegiatan pembelajaran harus berorientasi pada siswa. Pendekatan ini akan merangsang pengembangan kemampuan berfikir tingkat tinggi siswa, memberi keleluasaan pada siswa untuk belajar bekerja sama dan mendukung mereka untuk bersosialisasi terhadap apa saja yang ada dalam hidupnya).39

Heriberitus Sumarjo analyzes that KTSP is oriented in competence. It means the forms of purposes and contents will focus on master some competences like cognitive aspect, psychometric aspect, and affective aspect. If the students had get that competences, so their skill, their attitude and their good habitual will appear automatically.

There are five steps to implement the KTSP, they are:

- Standard Competence

There are some competencies or skills which are standardized for certain level, class, and semester.

---

39Educare, No. 4…KTSP, p. 5-6
• Competence Base
  There are certain skills to make some competencies

• Indicator
  There are consisting of specific competencies which are making as a guide or criteria in evaluation to decide a student in competence or not

• Material
  This is consist of everything will be student learn and comprehension about the process and result in order to measure the student’s comprehension about the competencies which are studied.

• Evaluation
  This is an activity to collect and use the information the process and result in order to measure the student’s competencies which are studied. 40

Drs. Samuel Bambang Muharyono H.S. stated that KTSP is operational curriculum which is arranged and developed by each education unit and he describes the contents of KTSP:
1. The purpose of education in the every level education unit
2. Describing vision and mission
3. Curriculum structure and time table of every subjects
4. Syllabus
5. Education calendar
6. The guide of service and function of educational facilities (laboratory and library)

He added six principles in developing KTSP:
1. Focus on student’s need, potency, progress, interest and also their environment
2. Manner and in the union
3. Response to the progress of knowledge, technology, and art
4. Relevance with the life needed
5. Long life education
6. Balance between national interest and region interest. 41

40Educare, No. 4,…, KTSP, p. 40
E. The Criteria of Conformity of Content Validity

In starting the level conformity of content validity, the writer uses the percentage formula as below:

\[ C = \frac{a}{b} \times 100\% \]

Notes:
- \( C \) = Conformity Level
- \( a \) = the indicators indicates which are found in the items of English summative test
- \( b \) = Indicators which are recommended in the KTSP SMA 87 English curriculum.

Then after the writer computes the percentage of conformity, the writer makes the conversion of score, which has same standard as Arikunto made:

- \( 76 – 100\% \) = Good
- \( 56 – 75\% \) = Sufficient
- \( 40 – 55\% \) = Less Good
- \(<40\% \) = Bad

---

\(^{41}\)Educare, No. 4...., KTSP, p. 26
\(^{42}\)Arikunto, Prosedur...., p. 313
CHAPTER III
THE IMPLEMENTATION OF THE RESEARCH

This chapter presents the research methodology which is divided into method of the study, time and location, technique of data collecting and technique of data analysis. Then, the findings which are divided into description of the data and interpretation of the data.

A. RESEARCH METHODOLOGY

1. Method of The Study

The method of this study is descriptive and evaluative analysis validity of English summative test made by the English teacher at SMAN 87 Jakarta. This study is aimed at describing and evaluating the content validity of the English summative test was designed by English teacher of SMAN 87 Jakarta.

The descriptive study is designed to obtain the current status of phenomena and is directed toward determining the nature of situation as it exists at the time of study that is in SMAN 87 Jakarta. It is called evaluate because it evaluates objectively the English summative test validity at odd semester at second grade students of SMAN 87 Jakarta based on the School-Level Curriculum (KTSP).

This study also called analysis because it analyzes objectively whether the material coverage in English summative test which is made by the English teacher of SMAN 87 Jakarta in line with the English syllabus and Guideline in English textbook assessment recommended by Pusat Kurikulum.
2. The Object of the Research

The object of the research is the test items of English Summative Test of SMA Negeri 87 Jakarta in the second semester 2008-2009 academic years, which consist of 40 items all of them are multiple choice. And the test held on Tuesday, 2nd of June 2009.

3. The Time and Location of Research

The research was administrated since June until July 2009, and location was at SMA Negeri 87 Jakarta at Jl. Mawar II Kel. Bintaro South Jakarta.

4. The Technique of Data Collecting

There are two instruments used to get the data represent two variables they are English summative test and English Syllabus. In more detailed, the writer performs the analysis as follows:
1. Exploring the English KTSP of SMA Negeri 87 Jakarta at the second grade.
2. Identifying the items of English Summative Test and classify them according to the functional skills.
3. Comparing the reality of the test with the English KTSP by making checklist.
4. Writing the findings.

Interpreting the data by calculating the total percentage of the conformity.

5. The Technique of Data Analysis

Technique of data analysis used descriptive analysis technique by using percentage; it means described using the table percentage. The conformity of the summative test items to the functional skills at syllabus. The writer used formula the conformity of the content validity as follows:
\[ C = \frac{a}{b} \times 100\% \]

Notes:

\( C \) = Conformity Level

\( a \) = the indicators indicates which are found in the items of English summative test

\( b \) = Indicators which are recommended in the KTSP SMA 87 English curriculum.

Then after the writer computes the percentage of conformity, the writer makes the conversion of score, which has same standard as Arikunto made:

- 76 – 100% = Good
- 56 – 75% = Sufficient
- 40 – 55% = Less Good
- <40% = Bad.1

B. RESEARCH FINDINGS

1. The Description of the Data

Here are the descriptions of the test seen from functional skills. the writer is able to describe the data as follows:

<table>
<thead>
<tr>
<th>No</th>
<th>Functional Skills</th>
<th>Item Numbers</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Expression of opinion</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Asking of Opinion</td>
<td>21, 39</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Expression of satisfaction</td>
<td>28, 33</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>Expression of unsatisfaction</td>
<td>38</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Advice</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>Warning</td>
<td>25, 34</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>Asking permission</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>Expression of Feel (relief, pain and pleasure)</td>
<td>26, 40</td>
<td>2</td>
</tr>
</tbody>
</table>

1Arikunto, *Prosedur*..., p. 313
2. The Analysis of Data

In this section, the writer analyzed the collected data by using qualitative and quantitative methods focusing on the conformity based on the functional skills.

First, he uses qualitative method to identify the item tests that conform only to the functional skills. Second, he uses quantitative method to calculate the total percentage of the conformity (see data interpretation).

The total item members of English Summative Test are 40 items all of them are multiple choice.

a. Functional Skills

Related to the functional skills, there are 20 items in the English Summative Test that measured the functional skills, so, the writer calculates that \( \frac{20}{40} \times 100\% = 50\% \), and they are distributed into 8 from 10 functional skills that must be covered.

Here is the 20 items test that measured 8 functional skills as suggested in the curriculum:

<table>
<thead>
<tr>
<th>No</th>
<th>Functional Skills</th>
<th>Item Numbers</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Attitude</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>Expression of Feeling (love, sadness, embarrassment, anger and annoyance)</td>
<td>23, 24, 29, 30, 31, 32, 35, 36, 37</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>10 Functional Skills</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

**Table 3.2**

The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus

<table>
<thead>
<tr>
<th>The Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>27. Henry: Rudy is doing fine. Rita: … He’s so rude.</td>
</tr>
<tr>
<td>a. I’m sure</td>
</tr>
<tr>
<td>b. I am with you.</td>
</tr>
<tr>
<td>c. I don’t think so</td>
</tr>
<tr>
<td>d. That’s my fault.</td>
</tr>
<tr>
<td>e. I have no resistance</td>
</tr>
</tbody>
</table>
The item number 27 above require the student’s ability to understand kind of “expression opinion”. So it can be understood that the item conform to the functional skill of “Expression of Opinion”.

### Table. 3.3
The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus

<table>
<thead>
<tr>
<th>The Question</th>
</tr>
</thead>
</table>
| 21. Indra : What do you think of the water transportation plan?  
Rika: …, it can decline the problem of traffic jam.  
a. It’s scary  
b. Worriging  
c. It’s alright  
d. I don’t care  
e. I have no idea |
| 39. Tika : Any objection about this idea?  
Rina: … It’s good.  
a. We have to make a review.  
b. It’s too complicated.  
c. I have no idea at all.  
d. We cannot make it.  
e. I can accept it. |

The items number 21 and 39 measures the students’ ability to choose asking the opinion, and it can be concluded that the items above conforms to functional skills of “Asking Opinion”

### Table. 3.4
The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus

<table>
<thead>
<tr>
<th>The Question</th>
</tr>
</thead>
</table>
| 28. Nancy : …  
Nanda : I just enjoy what I am doing.  
a. Very brilliant  
b. It sounds great.  
c. That’s wonderful  
d. You are very dedicated.  
e. How do you like your job? |
The Question

33. Ida : …
   Hendry : Thank you.
   a. I miss your plan
   b. Your plan is realistic.
   c. Your plan is so weird.
   d. Your plan will be refused.
   e. No one can understand your plan.

The items number 28 and 33 measures the students’ ability to choose the expression of satisfaction, and it can be concluded that the items above conforms to functional skills of “Expression of Satisfaction”

Table. 3.5
The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus

The Question

   Ica : … I really need the book today.
   a. You make me disappointed
   b. That’s OK. Bring it soon.
   c. Don’t tell it.
   d. Never mind
   e. Don’t worry

The item number 38 above require the student’s ability to understand kind of “expression of unsatisfied”. So it can be understood that the item conform to the functional skill of “Expression of Unsatisfaction”.
Table. 3.6
The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus

<table>
<thead>
<tr>
<th>The Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>25. John : ....</td>
</tr>
<tr>
<td>Ica : I’ll warn him</td>
</tr>
<tr>
<td>a. I know him</td>
</tr>
<tr>
<td>b. He’s so nice.</td>
</tr>
<tr>
<td>c. I can’t believe it.</td>
</tr>
<tr>
<td>d. He makes me mad</td>
</tr>
<tr>
<td>e. I realize about his wrong doings</td>
</tr>
<tr>
<td>34. Mother : …</td>
</tr>
<tr>
<td>Son : Yes, mom</td>
</tr>
<tr>
<td>a. I have no idea why you don’t eat lunch</td>
</tr>
<tr>
<td>b. Why don’t you have your lunch?</td>
</tr>
<tr>
<td>c. Have you got your lunch, darling</td>
</tr>
<tr>
<td>d. Where’s your lunch?</td>
</tr>
<tr>
<td>e. Is the lunch ready?</td>
</tr>
</tbody>
</table>

The items number 25 and 34 measures the students’ ability to choose the right about “Warning”, and it can be concluded that the items above conforms to functional skills of “Warning”

Table. 3.7
The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus

<table>
<thead>
<tr>
<th>The Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>26. Teacher : …</td>
</tr>
<tr>
<td>Rina : Forgive me, sir.</td>
</tr>
<tr>
<td>a. It was really unpleasant</td>
</tr>
<tr>
<td>b. I’m afraid it happened</td>
</tr>
<tr>
<td>c. That’s my fault.</td>
</tr>
<tr>
<td>d. I can solve it</td>
</tr>
<tr>
<td>e. I’ll do it.</td>
</tr>
</tbody>
</table>
### The Questions

<table>
<thead>
<tr>
<th>The Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>40.</strong> Indah : She’s very nice. …</td>
</tr>
<tr>
<td>Frida : I like her, too</td>
</tr>
<tr>
<td>a. I love him</td>
</tr>
<tr>
<td>b. No specialty</td>
</tr>
<tr>
<td>c. It’s depressing</td>
</tr>
<tr>
<td>d. He’s too clever</td>
</tr>
<tr>
<td>e. I can’t understand him</td>
</tr>
</tbody>
</table>

The items number 26 and 40 above measure the students’ ability to choose the right answer about “Expression of Feeling (relief, pain and pleasure)”, and it can be concluded that the items above conform to functional skills of “Expression of Feeling (relief, pain and pleasure)”.

### Table 3.8

**The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus**

<table>
<thead>
<tr>
<th>The Question</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>22.</strong> Father : Can you adjust your self, son?</td>
</tr>
<tr>
<td>Son : Yes, dad. ..</td>
</tr>
<tr>
<td>a. I miss you</td>
</tr>
<tr>
<td>b. How is life ?</td>
</tr>
<tr>
<td>c. I like the city.</td>
</tr>
<tr>
<td>d. I can stand on it.</td>
</tr>
<tr>
<td>e. Every thing is OK.</td>
</tr>
</tbody>
</table>

The item number 22 above require the student’s ability to understand kind of “expression of attitude”. So it can be understood that the item conform to the functional skill of “Expression of Attitude”.
Table. 3.9
The English summative test items used at SMAN 87 Jakarta which were conformed to the English Syllabus

<table>
<thead>
<tr>
<th>The Questions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>23. Ervan : I’m very sorry to hear the tragic accident of your nephew.</td>
<td></td>
</tr>
<tr>
<td>Reny : ...</td>
<td></td>
</tr>
<tr>
<td>a. I’m fine</td>
<td></td>
</tr>
<tr>
<td>b. I’m good</td>
<td></td>
</tr>
<tr>
<td>c. Oh. Thanks</td>
<td></td>
</tr>
<tr>
<td>d. I’m praised</td>
<td></td>
</tr>
<tr>
<td>e. That’s alright</td>
<td></td>
</tr>
<tr>
<td>24. Mira : You are wearing different socks.</td>
<td></td>
</tr>
<tr>
<td>Eky : ... How can I make a stupid thing.</td>
<td></td>
</tr>
<tr>
<td>a. What’s up?</td>
<td></td>
</tr>
<tr>
<td>b. I don’t care</td>
<td></td>
</tr>
<tr>
<td>c. I have no idea.</td>
<td></td>
</tr>
<tr>
<td>d. I’m not satisfied.</td>
<td></td>
</tr>
<tr>
<td>e. I am really ashamed.</td>
<td></td>
</tr>
<tr>
<td>29. Mr. Jaya : I lost my laptop.</td>
<td></td>
</tr>
<tr>
<td>Mr. Hendrawan: ...</td>
<td></td>
</tr>
<tr>
<td>a. I’m unhappy to hear that.</td>
<td></td>
</tr>
<tr>
<td>b. I greatly appreciate</td>
<td></td>
</tr>
<tr>
<td>c. Never mind</td>
<td></td>
</tr>
<tr>
<td>d. Is it okay?</td>
<td></td>
</tr>
<tr>
<td>e. It’s funny.</td>
<td></td>
</tr>
<tr>
<td>30. Ervan : Did you get a good score?</td>
<td></td>
</tr>
<tr>
<td>Indah : ... That’s not what I expect.</td>
<td></td>
</tr>
<tr>
<td>a. Certainly</td>
<td></td>
</tr>
<tr>
<td>b. I love it.</td>
<td></td>
</tr>
<tr>
<td>c. Shameful</td>
<td></td>
</tr>
<tr>
<td>d. I’m so proud.</td>
<td></td>
</tr>
<tr>
<td>e. You must know about it</td>
<td></td>
</tr>
<tr>
<td>31. Rina : It’s deafening. ...</td>
<td></td>
</tr>
<tr>
<td>Ruly : OK. I’ll tell them keep quiet.</td>
<td></td>
</tr>
<tr>
<td>a. I appreciate.</td>
<td></td>
</tr>
<tr>
<td>b. I’m stunned.</td>
<td></td>
</tr>
<tr>
<td>c. I’m irritated</td>
<td></td>
</tr>
<tr>
<td>d. They are mad</td>
<td></td>
</tr>
<tr>
<td>e. They can do it.</td>
<td></td>
</tr>
</tbody>
</table>
### The Questions

<table>
<thead>
<tr>
<th>Number</th>
<th>Dialogue</th>
<th>Choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>32.</td>
<td>Fery: ... I wish the music were not so loud.</td>
<td>a. It’s disappointing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b. It’s annoying</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c. It’s amazing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d. It’s lovely</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e. It’s odd</td>
</tr>
<tr>
<td></td>
<td>Fina: Sorry, I’ll turn it down.</td>
<td></td>
</tr>
<tr>
<td>35.</td>
<td>Dita: ... We lost.</td>
<td>a. It’s surprising</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b. It’s amazing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c. I feel sad</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d. I made it</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e. Terrific</td>
</tr>
<tr>
<td></td>
<td>Frida: You played well, though.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36.</td>
<td>Mira: Get in</td>
<td>a. I like it very much</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b. It’s shameful</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c. I’m pleased</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d. I’m trying</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e. I will do it</td>
</tr>
<tr>
<td></td>
<td>Rudy: ... I’m very late</td>
<td></td>
</tr>
<tr>
<td>37.</td>
<td>Indra: They have not paid for this month..</td>
<td>a. Cool</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b. Great</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c. Perfect</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d. Irritating</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e. extraordinary</td>
</tr>
<tr>
<td></td>
<td>Ika: They do it again. It’s ...</td>
<td></td>
</tr>
</tbody>
</table>

The items number 23, 24, 29, 30, 31, 32, 35, 36, 37 above need the student’s comprehension about Expression of Feeling (love, sadness, embarrassment, anger and annoyance). So it can be understood that the items conform to the functional skills of: “Expression of Feeling (Love, sadness, embarrassment, anger and annoyance)”.
3. Data Interpretation

Based on the data of items analysis result, we can see the table as follows:

**Table. 3.10**

The conformity question items based on the Functional Skills

<table>
<thead>
<tr>
<th>Conformity</th>
<th>Item Number</th>
<th>Curriculum Measured</th>
<th>Total Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>The conformity question items</td>
<td>21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40</td>
<td>8 of 10 Functional Skills suggested</td>
<td>80%</td>
</tr>
</tbody>
</table>

To interpret what the writer has already analyzed, he uses the criteria suggested by Arikunto:

**Table 3.11**

Criteria of Test Result Percentage

<table>
<thead>
<tr>
<th>Percentage Range</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>Less Good</td>
</tr>
<tr>
<td>&lt;40%</td>
<td>Bad</td>
</tr>
</tbody>
</table>

Viewed from the functional skills, the writer finds there are 20 items measuring eight of ten function skills to be measured. So the writer calculates the percentage of functional skills with the formula as follows:

\[ Cc = \frac{a}{b} \times 100\% \]

Notes:

C = The Level of Functional skills Conformity  
\( a \) = Functional skills found in the test items  
\( b \) = the Functional skills of the curriculum suggested  
\[ Cc = \frac{8}{10} \times 100\% = 80\% \]
And it fall in range of 76 – 100% which means “Good”

And having analyzed the result, the percentage of the total conformity is 80% of the functional skills. And it is in range of 76 – 80% which means “Good”.
RESEARCH FINDINGS

A. The Description of the Data

Here are the descriptions of the test seen from functional skills. the writer is able to describe the data as follows:

<table>
<thead>
<tr>
<th>No</th>
<th>Functional Skills</th>
<th>Item Numbers</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Expression of opinion</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Asking of Opinion</td>
<td>21, 39</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Expression of satisfaction</td>
<td>28, 33</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>Expression of unsatisfaction</td>
<td>38</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Advice</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Warning</td>
<td>25, 34</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>Asking permission</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Expression of Feel (relief, pain and pleasure)</td>
<td>26, 40</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>Attitude</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>Expression of Feeling (love, sadness, embarrassment, anger and annoyance)</td>
<td>23, 24, 29, 30, 31, 32, 35, 36, 37</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>10 Functional Skills</strong></td>
<td><strong>20</strong></td>
</tr>
</tbody>
</table>

B. The Analysis of Data

In this section, the writer analyzed the collected data by using qualitative and quantitative methods focusing on the conformity based the functional skills.

First, he uses qualitative method to identify the item tests that conform only to the functional skills. Second, he uses quantitative method to calculate the total percentage of the conformity (see data interpretation).

The total item members of English Summative Test are 40 items all of them are multiple choice.

1. Functional Skills
Related to the functional skills, there are 10 items in the English Summative Test that measured the functional skills, so, the writer calculates that ($\frac{20}{40} \times 100\% = 50\%$), and they are distributed into 8 from 10 functional skills that must be covered.

Here are the 20 items test that measured 8 functional skills as suggested in the curriculum:

<table>
<thead>
<tr>
<th>The Question</th>
</tr>
</thead>
</table>
| **27. Henry:** Rudy is doing fine.  
Rita: ... He’s so rude.  
 a. I’m sure  
 b. I am with you.  
 c. I don’t think so  
 d. That’s my fault.  
 e. I have no resistance |

The item number 27 above require the student’s ability to understand kind of “expression opinion”. So it can be understood that the item conform to the functional skill of “Expression of Opinion”.

<table>
<thead>
<tr>
<th>The Question</th>
</tr>
</thead>
</table>
| **21. Indra:** What do you think of the water transportation plan?  
Rika: ... , it can decline the problem of traffic jam.  
 a. It’s scary  
 b. Worrying  
 c. It’s alright  
 d. I don’t care  
 e. I have no idea |

<table>
<thead>
<tr>
<th>The Question</th>
</tr>
</thead>
</table>
| **39. Tika:** Any objection about this idea?  
Rina: ... It’s good.  
 a. We have to make a review.  
 b. It’s too complicated.  
 c. I have no idea at all. |
d. We cannot make it.
e. I can accept it.

The items number 21 and 39 measures the students’ ability to choose asking the opinion, and it can be concluded that the items above conforms to functional skills of “Asking Opinion”

The Question

28. Nancy : ...
   Nanda : I just enjoy what I am doing.
a. Very brilliant
b. It sounds great.
c. That’s wonderful
d. You are very dedicated.
e. How do you like your job?

33. Ida : ...
Hendry : Thank you.
a. I miss your plan
b. Your plan is realistic.
c. Your plan is so weird.
d. Your plan will be refused.
e. No one can understand your plan.
The items number 28 and 33 measures the students’ ability to choose the expression of satisfaction, and it can be concluded that the items above conforms to functional skills of “Expression of Satisfaction”

<table>
<thead>
<tr>
<th>The Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ica : … I really need the book today.</td>
</tr>
<tr>
<td>a. You make me disappointed</td>
</tr>
<tr>
<td>b. That’s OK. Bring it soon.</td>
</tr>
<tr>
<td>c. Don’t tell it.</td>
</tr>
<tr>
<td>d. Never mind</td>
</tr>
<tr>
<td>e. Don’t worry</td>
</tr>
</tbody>
</table>

The item number 38 above require the student’s ability to understand kind of “expression of unsatisfied”. So it can be understood that the item conform to the functional skill of “Expression of Unsatisfaction”.

<table>
<thead>
<tr>
<th>The Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>25. John : …</td>
</tr>
<tr>
<td>Ica : I’ll warn him</td>
</tr>
<tr>
<td>a. I know him</td>
</tr>
<tr>
<td>b. He’s so nice.</td>
</tr>
<tr>
<td>c. I can’t believe it.</td>
</tr>
<tr>
<td>d. He makes my mad</td>
</tr>
<tr>
<td>e. I realize about his wrong doings</td>
</tr>
</tbody>
</table>

34. Mother : …
Son : Yes, mom
   a. I have no idea why you don’t eat lunch
   b. Why don’t you have your lunch?
   c. Have you got your lunch, darling
   d. Where’s your lunch?
e. Is the lunch ready?

The items number 25 and 34 measures the students’ ability to choose the right about “Warning”, and it can be concluded that the items above conforms to functional skills of “Warning”.

<table>
<thead>
<tr>
<th>The Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>26. Teacher : …</td>
</tr>
<tr>
<td>Rina : Forgive me, sir.</td>
</tr>
<tr>
<td>a. It was really unpleasant</td>
</tr>
<tr>
<td>b. I’m afraid it happened</td>
</tr>
<tr>
<td>c. That’s my fault.</td>
</tr>
<tr>
<td>d. I can solve it</td>
</tr>
<tr>
<td>e. I’ll do it.</td>
</tr>
</tbody>
</table>

40. Indah : She’s very nice. …
Frida : I like her, too
a. I love him |
| b. No specialty |
| c. It’s depressing |
| d. He’s too clever |
| e. I can’t understand him |

The items number 26 and 40 above measure the students’ ability to choose the right answer about “Expression of Feeling (relief, pain and pleasure)”, and it can be concluded that the items above conform to functional skills of “Expression of Feeling (relief, pain and pleasure)”.

<table>
<thead>
<tr>
<th>The Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>22. Father : Can you adjust your self, son?</td>
</tr>
<tr>
<td>Son : Yes, dad. ..</td>
</tr>
<tr>
<td>a. I miss you</td>
</tr>
<tr>
<td>b. How is life ?</td>
</tr>
</tbody>
</table>
c. I like the city.
d. I can stand on it.
e. Every thing is OK.

The item number 22 above require the student’s ability to understand kind of “expression of attitude”. So it can be understood that the item conform to the functional skill of “Expression of Attitude”.

### The Questions

#### 23. Ervan : I’m very sorry to hear the tragic accident of your nephew.
Reny : …

- a. I’m fine
- b. I’m good
- c. Oh. Thanks
- d. I’m praised
- e. That’s alright

#### 24. Mira : You are wearing different socks.
Eky : … How can I make a stupid thing.

- a. What’s up?
- b. I don’t care
- c. I have no idea.
- d. I’m not satisfied.
- e. I am really ashamed.

#### 29. Mr. Jaya : I lost my laptop.
Mr. Hendrawan: …

- a. I’m unhappy to hear that.
- b. I greatly appreciate
- c. Never mind
- d. Is it okay?
- e. It’s funny.
30. Ervan : Did you get a good score?
   Indah : … That’s not what I expect.
   a. Certainly
   b. I love it.
   c. Shameful
   d. I’m so proud.
   e. You must know about it

31. Rina : It’s deafening. …
   Ruly : OK. I’ll tell them keep quiet.
   a. I appreciate.
   b. I’m stunned.
   c. I’m irritated
   d. They are mad
   e. They can do it.

32. Fery : …I wish the music were not so loud.
   Fina : Sorry. I’ll turn it down.
   a. It’s disappointing
   b. It’s annoying
   c. It’s amazing.
   d. It’s lovely.
   e. It’s odd.

35. Dita : … We lost.
   Frida : You played well, though.
   a. It’s surprising
   b. It’s amazing
   c. I feel sad.
   d. I made it.
   e. Terrific
36. Mira : Get in
Rudy : … I’m very late
   a. I like it very much
   b. It’s shameful
   c. I’m pleased
   d. I’m trying.
   e. I will do it

37. Indra : They have not paid for this month..
Ika    : They do it again. It’s …
   a. Cool
   b. Great
   c. Perfect
   d. Irritating
   e. extraordinary

The items number 23, 24, 29, 30, 31, 32, 35, 36, 37 above need the student’s comprehension about Expression of Feeling (love, sadness, embarrassment, anger and annoyance). So it can be understood that the items conform to the functional skills of “Expression of Feeling (Love, sadness, embarrassment, anger and annoyance”).

C. Data Interpretation

Based on the data of items analysis result, we can see the table as follows:

<table>
<thead>
<tr>
<th>Conformity</th>
<th>Item Number</th>
<th>Curriculum Measured</th>
<th>Total Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>The conformity question items based on</td>
<td>21, 22, 23, 24, 25,</td>
<td>8 of 10</td>
<td>80%</td>
</tr>
<tr>
<td>the Functional Skills</td>
<td>26, 27, 28, 29, 30,</td>
<td>Functional Skills</td>
<td></td>
</tr>
<tr>
<td></td>
<td>31, 32, 33, 34, 35,</td>
<td>suggested</td>
<td></td>
</tr>
<tr>
<td></td>
<td>36, 37, 38, 39, 40</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
To interpret what the writer has already analyzed, he uses the criteria suggested by Arikunto:

Table
Criteria of Test Result Percentage

<table>
<thead>
<tr>
<th>Percentage Range</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>Less Good</td>
</tr>
<tr>
<td>&lt;40%</td>
<td>Bad</td>
</tr>
</tbody>
</table>

Viewed from the functional skills, the writer finds there are 20 items measuring eight of ten function skills to be measured. So the writer calculates the percentage of functional skills with the formula as follows:

\[ Cc = \frac{a}{b} \times 100\% \]

Notes:

- \( C \) = The Level of Functional skills Conformity
- \( a \) = Functional skills found in the test items
- \( b \) = the Functional skills of the curriculum suggested

\[ Cc = \frac{8}{10} \times 100\% = 80\% \]

And it fall in range of 76 – 100% which means “**Good**”

And having analyzed the result, the percentage of the total conformity only 80% of the functional skills. And it is in range of 76 – 80% which means “**Good**”.
CHAPTER IV
CONCLUSION AND SUGGESTION

This chapter presents the conclusion and suggestion from the researcher based on the research done at SMA Negeri 87 Jakarta.

A. Conclusion

Based on the data analysis and the interpretation in the previous chapter, the writer would like to elaborate the conformity of English Summative Test of the odd semester for the second grade of students of SMA Negeri 87 Jakarta to their English KTSP.

Of the total items contained by English Summative Test of the odd semester for the second grade of SMA Negeri 87 Jakarta to their English KTSP there are 10 items measuring eight of ten function skills to be measured \((8/10 \times 100\% = 80\%)\). And it falls in range of 76 – 100% which means “Good”.

And having analyzed the result, percentage of the total conformity is 80% of the functional skills. And it is in the area of 76 – 100% which means “GOOD”.

B. Suggestions

For the conclusion written above, the writer would like to give some suggestions as follows:

1. The test items should be conformity to the curriculum category used; the items are fulfilling to the curriculum.
2. To fulfill the characteristics of a good test, the item numbers and each indicators of English KTSP must be examined and be analyzed before.
3. The teacher should revise invalid items, so they can be used for the next evaluation.

4. The school has to keep more opportunity to develop their competence in making the test.
BIBLIOGRAPHY


Definition of assessment at Pittsburg State University, this article accessed on September 13th 2009 from [http://www.diknas.com](http://www.diknas.com)


Puskur-Dit PLP/Wisata/1 September 2003


Sudijono, Anas, Drs., Prof., *Pengantar Evaluasi Pendidikan*, Jakarta: PT. Raja Grafindo Persada, 2005