

**THE QUALITY OF ENGLISH SUMMATIVE TEST ITEM  
OF MULTIPLE CHOICE MADE BY THE ENGLISH  
TEACHERS OF MTsN-1 MODEL  
PALANGKA RAYA**

**THESIS**

Presented to the Department of Education of the State Islamic College of  
Palangka Raya in Partial Fulfillment of the Requirements for  
the Degree of Sarjana Pendidikan Islam



**By:**

**NOOR AMALIA SHANTI**  
**SRN. 0301120110**

**THE ISLAMIC STATE COLLEGE OF PALANGKA RAYA  
THE DEPARTMENT OF EDUCATION  
THE STUDY PROGRAM OF ENGLISH EDUCATION  
2009**

Palangka Raya, September 2009

**OFFICIAL NOTE**

Case: Please Examined

To: Director of the Islamic State College of  
Palangka Raya  
in  
Palangka Raya

Peace be unto you and God's mercy and blessing as well.

By reading and analyzing of your thesis's revision. We think that the thesis in the name  
of:

**Name : NOOR AMALIA SHANTI**

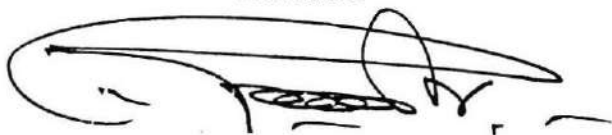
**SRN : 0301120110**

**Title : THE QUALITY OF ENGLISH SUMMATIVE TEST ITEM OF  
MULTIPLE CHOICE MADE BY THE ENGLISH TEACHERS OF  
MTsN-1 MODEL PALANGKA RAYA**

It can be examined (munaqasah) to get the degree of Sarjana Pendidikan Islam (S.Pd.I) in  
English Education of the Department of Education STAIN Palangka Raya.  
Therefore it has been expected that this thesis can be examined (munaqasah) soon.

*Wassalamualaiikum Wr. Wb.*

Advisor I



**Dr. H. KHAIRIL ANWAR, M.Ag.**  
**NIP.19630118 1991031 002**

Advisor II,



**Dra. HALIMAH, M.Pd.**  
**NIP. 196712326 199603 2 003**

## APPROVAL OF THE THESIS ADVISORY COMMITTEE

Title of the Thesis : **THE QUALITY OF ENGLISH SUMMATIVE TEST ITEM OF  
MULTIPLE CHOICE MADE BY THE ENGLISH TEACHERS OF  
MTsN-1 MODEL PALANGKA RAYA**

Name : Noor Amalia Shanti  
SRN : 0301120110  
Department : Education  
Study Program : English Education  
Level : (S-1)

Palangka Raya , September 2009

Approved by

Advisor I



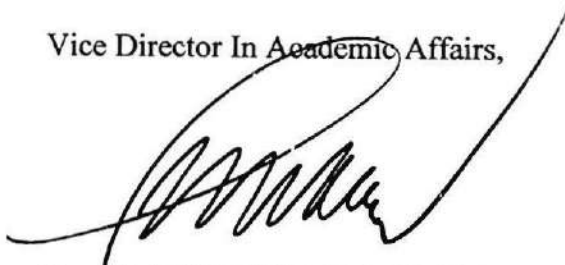
**Dr. H. KHAIRIL ANWAR, M.Ag.**  
NIP. 19630118 199103 1 002

Advisor II,



**Dra. HALIMAH, M.Pd.**  
NIP. 19671226 199603 2 003

Vice Director In Academic Affairs,



**Drs. H. ABU BAKAR H. M., M.Ag.**  
NIP. 19551231 198303 1 026

The Chair of Education Department,



**Hj. HAMIDAH M.A.**  
NIP. 19700425 199703 2 003

## LEGALIZATION OF THE THESIS EXAMINING COMMITTEE

This thesis entitles “ **THE QUALITY OF ENGLISH SUMMATIVE TEST ITEM OF MULTIPLE CHOICE MADE BY THE ENGLISH TEACHERS OF MTsN-1 MODEL PALANGKA RAYA**” In the name of Noor Amalia Shanti, and her Student Registration Number is 0301120110. It has been examined in Examining Team of the State Islamic College of Palangka Raya on :

Day : Thursday

Date : Dzulqa'dah 15<sup>th</sup> 1430 H November, 19<sup>th</sup> 2009

Time : 15.00 – 16.30

Palangka Raya, 19<sup>th</sup> November 2009

Team Examiners

1. Drs. H. Abdul Qodir M.Pd.  
Chairman / member

(.....)

2. Ristati M.Hum.  
Member

(.....)

3. Dr. H. Khairil Anwar, M.Ag.  
Member

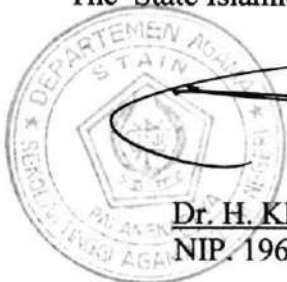
(.....)

4. Dra. Halimah M.Pd.  
Secretary / Member

(.....)

The State Islamic College of Palangka Raya

Director,



Dr. H. Khairil Anwar, M.Ag.  
NIP. 19630118 1991031 002



## ACKNOWLEDGEMENT

*Alhamdulillahirobbil'alamiin.* In the name of Allah, the most Graceful and Merciful. Thanks to Allah the lord of the world who has given guidance, healthy, ability till writer could finish this thesis entitled: the Quality of English Summative Test items of Multiple Choice made by the English teachers of MTsN-1 Palangka Raya.

The writer would like to thanks for all people who has given supports, advices and opportunities during the whole process so writer could finish this thesis. Therefore, the writer would like to thanks to:

1. Dr. H. Khairil Anwar, M.Ag., the Director of the Islamic State College of Palangka Raya and also as my first advisor for his guidance and encouragements to finish this thesis.
2. Hj. Hamidah, M.A., the Head of Tarbiyah Department, for her permission and also her cooperation so that the writer can complete the requirements to finish the thesis.
3. Dra. Halimah, M.Pd., as second advisor for her guidance, supports, and encouragement till the end of the writing this thesis.
4. Islamic State College of Palangka Raya, the entire components and staffs. And also, all the lectures of the English Education department for their supports, advices, and guidance during her study at Islamic State College of Palangka Raya till she can finish her study.

Finally, the writer would like to thanks to her loyal family and friends who always give her support, enthusiasms, attention and inspiration to finish her duty. May Allah show His blessing and mercies in every moment of our life.

NOOR AMALIA SHANTI

## DECLARATION OF AUTHENTICATION

In the name of God

I myself make declaration that this thesis entitles " **THE QUALITY OF ENGLISH SUMMATIVE TEST ITEM OF MULTIPLE CHOICE MADE BY THE ENGLISH TEACHERS OF MTsN-1 MODEL PALANGKA RAYA** " is truly my own writing. If it is not my own writing so, if given a citation and shown in the list of references.

If my own declaration is not right in this thesis so, I am ready to be given academic sanction namely, the cancellation of the degree of this thesis.

Palangka Raya, November, 2009

My Own Declaration



**NOOR AMALIA SHANTI**

SRN. 030 112 0110

***THE QUALITY OF ENGLISH SUMMATIVE TEST ITEM OF MULTIPLE CHOICE  
MADE BY THE ENGLISH TEACHERS OF MTsN-1 MODEL  
PALANGKA RAYA***

**ABSTRACT**

The problems of the study are, (1) Is the English summative test item of Multiple Choice made by the English teachers of MTsN-1 Model Palangka Raya valid? (2) Is the English summative test item of Multiple Choice made by the English teachers of MTsN-1 Model Palangka Raya reliable? (3) What is the discriminating power of English summative test items of Multiple Choice made by the English teachers of MTsN-1 Model Palangka Raya?

The objectives of the study are, (1) to measure the validity of the English summative test items of Multiple Choice made by the English teachers of MTsN-1 Model Palangka Raya, (2) to measure the reliability of the English summative test items of Multiple Choice made by the English teachers of MTsN Model-1 Palangka Raya, (3) to measure the discriminating power of English summative test items of Multiple Choice made by the English teachers of MTsN-1 Model Palangka Raya and its became the sample of the study. The item that was studied is English Multiple Choice item which consist 40 items

In this study the writer used quantitative method. The writer took the English summative test items of Multiple Choice made by the English teacher of MTsN-1 Model Palangka Raya for the second grade student at the first semester in academic year 2008/2009. The technique of collecting data was documentation. In technique of data analysis, the writer used some procedures. They are: (1) to arrange the students' score from the highest score to the lower score, (2) to identify the highest group and a lower group; and (3) to measure the validity, reliability and the discriminating power of item.

The result of the study showed that the content validity of the English summative test of Multiple Choice made by the English teacher of MTsN-1 Model Palangka Raya for the second grade student at the first semester in academic year 2008/2009 had fair validity. The reliability of the test was 0.75 so it was high reliability. And the result of discriminating power only 10 items (25%) so it was not discriminable.

**KUALITAS BUTIR SOAL SUMMATIF BAHASA INGGRIS BUATAN GURU  
YANG BERBENTUK PILIHAN GANDA PADA MTsN-1 MODEL  
PALANGKA RAYA**

**ABSTRAK**

Rumusan masalah dari penelitian ini adalah (1) Apakah soal summatif bahasa Inggris buatan guru yang berbentuk pilihan ganda valid? (2) Apakah soal tes summatif bahasa Inggris buatan guru yang berbentuk pilihan ganda reliabel? (3) Apa daya pembeda pada soal summatif bahasa Inggris buatan guru yang berbentuk pilihan ganda?

Tujuan dari penelitian ini adalah untuk mengukur kevalidan (konten), kereliabilitasan, dan daya pembeda dalam soal summatif bahasa Inggris buatan guru yang berbentuk pilihan ganda pada sekolah MTsN-1 Model Palangka Raya dan ini menjadi sampel dalam penelitian ini. Soal tes summatif bahasa Inggris pilihan ganda yang diteliti berjumlah 40 soal

Dalam penelitian ini penulis menggunakan metode kuantitatif karena menjelaskan hasilnya dalam bentuk angka (hitungan). Penulis meneliti soal summatif bahasa Inggris buatan guru yang berbentuk pilihan ganda pada sekolah MTsN-1 Model Palangka Raya di kelas VIII pada semester pertama tahun ajaran 2008/2009. Teknik pengumpulan data yang telah digunakan adalah dokumentasi. Pada teknik proses menganalisis data, penulis menggunakan beberapa prosedur, yaitu: (1) Menyusun skor siswa dari yang tertinggi sampai yang terendah, (2) Mengidentifikasi grup tertinggi dan grup terendah; dan (3) Mengukur kevalidan, kereliabilitasan dan daya pembeda pada soal.

Hasil penelitian di MTsN-1 Palangka Raya menggambarkan bahwa soal summatif bahasa Inggris buatan guru yang berbentuk pilihan ganda pada sekolah MTsN-1 Model Palangka Raya di kelas VIII pada semester pertama tahun ajaran 2008/2009 kurang valid. Kereliabilitasan soal 0,75 jadi dalam hal ini soal termasuk reliabel. Dan hanya 10 soal yang bisa dijadikan daya pembeda atau sekitar 25%.

## DEDICATION

*This thesis is dedicated to:*

*\*My beloved parents, Drs. H. Ahmad Kusasi and Masliana who have given me*

*love, affection, and lots of support;*

*\*My beloved sisters and brothers who always motivate me to finish this study;*

*\*My Aunt and My Uncle, Thanks you for your help;*

*\*The Coordinator of the English Education Study Program;*

*\*Dr. H. Khairil Anwar, M.Ag., as my fist advisor and Dra. Halimah, M.Pd., as my*

*second advisor who have guided and advised me during the thesis done.*

*\*My beloved friends, Moeza, Odah, DuLz, Thyeka, The2d, Bolz, Ainah, N'dank,*

*Figo, Ole, Narto, Rety, Eka, Ati, Lia, Imah n'dut and Ayuy who always*

*encourage and motivate me till I can finish my study.*

*May Allah always love and protect us in all of our ways  
till the end of the day.*

## MOTTO

... قُلْ هَلْ يَسْتَوِي الَّذِينَ يَعْلَمُونَ وَالَّذِينَ لَا يَعْلَمُونَ <sup>قُلْ</sup> إِنَّمَا يَتَذَكَّرُ <sup>قُلْ</sup>

أُولُوا الْأَلْبَابِ ﴿٩﴾

**“... say: Are those equal, those who know and those who do not know? It is those who are endued with understanding that receive admonition”.**

**(Q.S. Az-‘Zumar: 9)**

## TABLE OF CONTENTS

	Page
COVER PAGE .....	i
OFFICIAL NOTE .....	ii
APPROVAL OF THESIS ADVISORY COMMITTEE.....	iii
LEGALIZATION OF THE THESIS EXAMINING COMMITTEE.....	iv
ACKNOWLEDGMENTS .....	v
DECLARATION OF AUTHENTICATION .....	vi
ABSTRACT .....	vii
DEDICATION .....	viii
MOTTO .....	ix
TABLE OF CONTENTS .....	x
LIST OF TABLES .....	xii
LIST OF CHARTS .....	xiii
CHAPTER I INTRODUCTION	
A. Background of the Study .....	1
B. Problems of the Study .....	4
C. Objectives of the Study .....	4
D. Significances of the Study .....	5
E. Definition of concept .....	5
F. Delimitation of the study .....	7
G. Variable of the Study .....	7
H. Research Methodology .....	8
1. Method of the Study .....	8
2. The Subject of Study .....	9
3. The Technique of Collecting Data .....	9
5. The Technique of Processing Data .....	16
6. The Technique of Data Analysis Procedures.....	18
I. Framework of the Discussion .....	23

CHAPTER II	REVIEW OF RELATED LITERATURE	
A.	Test .....	24
B.	Uses and Function of Test .....	25
C.	Kind of Test .....	27
1.	Based on organization .....	27
2.	Based on Method of teaching .....	27
3.	Based on its Function .....	29
4.	Based on the Form of the Test .....	29
5.	Based on the Scoring system .....	30
D.	The principles in writing Multiple Choice .....	34
E.	The Strengths and The weaknesses .....	37
F.	The Characteristics of Good Test .....	42
1.	Validity .....	43
2.	Reliability .....	45
3.	Discriminating Power .....	46
G.	Test Procedure .....	48
H.	Scoring Test.....	48
I.	The Review of Research Finding .....	49
CHAPTER III	RESULT OF THE STUDY	
A.	The English Summative Test in MTsN-1 Model Palangka Raya.....	50
B.	The validity of English Summative test items of Multiple Choice made by the English teacher of MTsN- 1 Palangka Raya .....	50
C.	The reliability of English Summative test items of Multiple Choice made by the English teacher of MTsN- 1 Palangka Raya .....	52
D.	The discriminating power of English Summative test items of Multiple Choice made by the English teacher of MTsN-1 Palangka Raya.....	55
E.	The students' score of MTsN-1 Model in Summative test in academic year 2008/2009 .....	58
CHAPTER IV	CLOSING	
A.	Conclusion .....	60
B.	Suggestion .....	61
REFERENCES		
APPENDIX		



## LIST OF TABLES

Table 1	Content Specification of test items .....	13
Table 2	The Percentage of discriminating Power .....	56
Table 3	The distribution of Frequency and Percentage of students' scores in Summative test .....	58

### **LIST OF CHART**

Chart 1	The frequency and percentage of discriminating power ..	57
Chart 2	The percentage of students' scores in summative test....	59

# CHAPTER I

## CHAPTER I

### INTRODUCTION

#### A. Background of the study

Evaluation is one of the important processes in educational field. In teaching learning process, evaluation is very important to many facets of school program and for the teachers it's very important to measure the students' progress. After the educational program has begun, information on students' behavior is needed as to judge the success of the students in reaching the goals. The evaluation students' progress is a major aspect of the teacher's job. By the evaluation, the teacher will know how far the students understand about the materials that have been given, and the teacher can find out the students' weaknesses or difficulties during the teaching learning process.

There are some definitions of evaluation. According to Gronlund in Tinambunan, "Evaluation is the systematic process of collecting or analyzing and interpreting information to determine the extent to which pupils are achieving instructional objectives".<sup>1</sup> While Tyler in Arikunto states "*Evaluasi merupakan sebuah proses pengumpulan data untuk menentukan sejauh mana, dalam hal apa, dan bagaimana tujuan pendidikan sudah tercapai*".<sup>2</sup> (Evaluation is a process of gathering the data to decide how far, in what matter, and how the purpose of education has reached).

---

<sup>1</sup>Wilmar Tinambunan, *Evaluation of Students' Achievement*, Jakarta: Departemen Pendidikan dan Kebudayaan, 1988, p. 4.

<sup>2</sup>Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, Jakarta: Bumi Aksara, 2003, p. 3.

definition of evaluation is:

*“Evaluation is the systematic acquisition and assessment of information to provide useful feedback about some object”.*<sup>3</sup>

Both definitions agree that evaluation is a systematic endeavor and both use the deliberately ambiguous term 'object' which could refer to a program, policy, technology, person, need, activity, and so on. The latter definition emphasizes acquiring and assessing information rather than assessing worth or merit because all evaluation work involves collecting and sifting through data, making judgments about the validity of the information and of inferences we derive from it, whether or not an assessment of worth or merit results.

In teaching learning process to evaluate students' progress the teacher uses a test as an instrument. By the test the teacher can monitor students learning and diagnose strengths and weaknesses their students. According to Tuckman in Melya states that “a test is a device for sampling behavior or performance related to the skills, competencies, attitudes, or other characteristics of people”.<sup>4</sup> And Heaton says that “testing and teaching are so closely interrelated that is virtually impossible to work on either field without being concerned with the other”.<sup>5</sup>

---

<sup>3</sup><http://www.socialresearchmethods.net/kb/intreval.php> (On line April, 26<sup>th</sup> 2009).

<sup>4</sup>Melya Oktarini Y.B., “The Quality of English Summative Test Made by The English Teacher of SMP Kristen Palangka Raya”, *Skripsi*, Palangka Raya: Universitas Palangka Raya, 2007, p. 1, t.d.

<sup>5</sup>J.B Heaton, *Writing English Language Test*, London: Longman Group Ltd, 1979, p. 1.

It means both testing and teaching must be interrelated. Therefore, teachers must be good in teaching and constructing a good test as well. A good test is a basic thing that is really needed in evaluating the teaching learning process.

In evaluation, there are many kinds of the test. One of them is Multiple Choice that usually often used by the English teacher. It is because it can be more effective to measure many of the simple learning outcomes measured by the short item or completion, the true-false item, and the matching item. And it is flexible, plus the higher quality items are usually found in the Multiple Choice items. So far teachers in MTSN-1 Model always choose Multiple Choice for the summative test. They make their test based on syllabus's book

In this case, the writer focuses to the test items of Multiple Choice. It is caused the authenticity of Multiple Choice questions can be maximized by framing each question item as a case example. A half or even for the most education institution the achievement test is held by using Multiple Choice item. Even in a big scale activity, such as *Ujian Nasional* (National examination) and *Ujian Sekolah* (School examination). So, in this study the writer proposed a title is: **“THE QUALITY OF ENGLISH SUMMATIVE TEST ITEM OF MULTIPLE CHOICE MADE BY THE ENGLISH TEACHERS OF MTsN-1 MODEL PALANGKA RAYA”**. Through this study, the writer wants to know that the summative test items of Multiple Choice made by the English teachers of MTsN-1 Model Palangka Raya have fulfilled the criteria of a good test.

As it is known, MTsN-1 Palangka Raya is one of the junior high school which has Islamic characteristic. And it is a model school. So, as a model school it can be a good model for other schools, especially in developing to make a good test.

The writer also finds that there is one previous study, which almost similar to her study that is the quality of English summative test made by the English teacher of SMP Kristen Palangka Raya by Melya Oktarini Y.B. although the two studies talked about the same thing that is about the quality of English test, but actually the two studies focused on the different subject and object of the study.

### **B. Problems of the Study**

Based on the statement above, so it can be developed into the following problems:

1. Is the English summative test item of Multiple Choice made by the English teachers of MTsN-1 Model Palangka Raya valid?
2. Is the English summative test item of Multiple Choice made by the English teachers of MTsN-1 Model Palangka Raya reliable?
3. What is the discriminating power of English summative test items of Multiple Choice made by the English teachers of MTsN-1 Model Palangka Raya?

### **C. Objectives of the Study**

The objectives of this study are:

1. To measure the validity of the English summative test items of Multiple Choice made by the English teachers of MTsN-1 Model Palangka Raya.

2. To measure the reliability of the English summative test items of Multiple Choice made by the English teachers of MTsN Model-1 Palangka Raya.
3. To measure the discriminating power of English summative test items of Multiple Choice made by the English teachers of MTsN-1 Model Palangka Raya.

#### **D. Significances of the Study**

Theoretically, this study is expected to be beneficial to increase the English teacher in constructing a good test to measure student's achievement. Practically, this study gives the description of how valid, reliable, and the discriminating power of the Multiple Choice item made by the English teachers. It provides important information about how to construct the test based on the criteria of a good test and also purposed to share some contribution as source of information for further researches to conduct similar studies and the library of the Islamic State College of Palangka Raya.

#### **E. Definition of the Concept**

The concept here is related to the quality of English Summative test items of Multiple Choice made by the English teachers at the eighth grade students in the first semester of MTsN-1 Model Palangka Raya .

1. Quality is the standard of something when compared to other things like it; how good and bad something is.<sup>6</sup> From the meaning of quality before, the writer wants to measure about the quality. In this study, the quality that will be measured

---

<sup>6</sup>A.S.Hornby, *Oxford Advanced Learner's Dictionary of Current English*, Jonathan Crowther, Oxford: Oxford University Press, 1995, p. 950.



is English test items of Multiple Choice made by the English teacher of MTsN-1 Model Palangka Raya.

2. Summative test is a test that done by the teacher to measure the students ability at the end semester. Summative assessment is the formal testing of what has been learned in order to produce marks or grades which may be used for reports of various types.<sup>7</sup> In this case, the English summative test is made by the English teachers of MTsN-1 Model Palangka Raya.

3. Multiple Choice test is made up of items which present two or more responses, only one of them is correct of definitely better than the others.<sup>8</sup> An advantage of Multiple choice items is that they are quickly and objectively scored.

#### 4. Teacher Made Test

According to Slameto is :

*“tes buatan guru yaitu tes yang dibuat oleh guru untuk keperluan penilaian guru tersebut terhadap siswanya. Tes ini terutama tes hasil belajar dan biasanya berlaku untuk satu sekolah, bahkan kadang-kadang hanya untuk satu kelas saja”.*<sup>9</sup> ( test made by the teacher is a test that made by the teacher for assessing the student. This test especially in test result of learning and usually valid for one school, even sometimes for only one class).

5. Criteria of a good test, the writer only take some criteria of a good test like validity, reliability and discriminating power to measure whether the test is good or not.

---

<sup>7</sup>[http://www.highlandschools-virtualib.org.uk/lt/whole\\_learner/summative.htm](http://www.highlandschools-virtualib.org.uk/lt/whole_learner/summative.htm) (On line April 18th 2009).

<sup>8</sup>E.N. Gronlund, *Measurement and Evaluation in Teaching (fourth edition)*, New York: Macmillan Publishing Co., Inc. 1981, p. 221.

<sup>9</sup>Slameto, *Evaluasi Pendidikan*. Cet.2. Jakarta: Sinar Grafika Offset, 2001, p. 32.

## **F. Delimitation of the Study**

The study is only done to measure the validity, reliability and discriminating power of English Summative test items of Multiple Choice made by the English teachers of MTsN-1 Model Palangka Raya. The subject of this study is the English Summative test items of Multiple Choice made by the English teachers of MTsN-1 Model Palangka Raya to the eighth grade students in the first semester in academic year 2008/2009.

## **G. Variable of the Study**

Variable is *“gejala yang menjadi fokus peneliti untuk diamati. Variable itu sebagai atribut dari sekelompok orang atau obyek yang mempunyai variasi antara satu dengan yang lainnya dalam sekelompok itu”*.<sup>10</sup> ( indication that have been researched by researched. Variable as the attribute of a group of people or the object that have various among them).

This variable of this study is the features of the quality of the English test item of Multiple Choice made by the English teachers of MTsN-1 Model Palangka Raya.

Which consist the following:

1. Content and construct validity of the English test.
2. Reliability of the English test.
3. Discriminating power of the English test.

---

<sup>10</sup>Sugiyono, *Statistika Untuk Penelitian* , Bandung: Alfabeta, 2004, p. 2.

## H. Research Methodology

### 1. Time and Place of the Study

The place of this study held in MTsN-1 Model Palangka Raya since April 4<sup>th</sup> until June 4<sup>th</sup>, 2009 in the first semester in academic year 2008/2009.

### 2. Method of the Study

In this study, the writer used a quantitative method. According to Kountur, *“data kuantitatif adalah data yang dapat diukur sehingga dapat menggunakan statistik dalam pengujiannya”*.<sup>11</sup> (quantitative data is a data that it can be measured, so it can use statistic in its testing). Sukardi states that the writer efforts to describe the research activity on the certain object systematically and clearly in quantitative descriptive approach.

According to Arikunto state:

*“Penelitian Deskriptif merupakan penelitian yang dimaksudkan untuk mengumpulkan informasi mengenai status suatu gejala yang ada, yaitu keadaan gejala menurut apa adanya pada saat penelitian dilakukan. Penelitian deskriptif tidak memerlukan administrasi atau pengontrolan terhadap suatu perlakuan.”*<sup>12</sup> (Descriptive study intends to gather information about the status of the actual fact of tendency and it doesn't need any administration or controlling of action”).

In quantitative research the aim is to determine the relationship between one thing (an independent variable) and another (a dependent or outcome variable) in a population. Quantitative research designs are either descriptive (subjects usually measured once) or experimental (subjects measured before and

---

<sup>11</sup>Ronny Kountur, *Metode Penelitian*, Jakarta: Penerbit PPM, 2005, p. 16.

<sup>12</sup>Suharsimi Arikunto, *Manajemen Penelitian*. Jakarta: Rineka Cipta. 2000. p. 309.

after a treatment). A descriptive study establishes only associations between variables. An experiment establishes causality.

In this way, quantitative method means to find out the data and describe the phenomena through the process being observed in detail and clearly as reality that found in English Summative test items of Multiple Choice made by the English teachers of MTsN-1 Model Palangka Raya.

### **3. The Subject of the Study**

The subject of this study is the English summative test items of Multiple Choice. The Numbers of the subject are 40 items. The writer takes all of the English summative test of Multiple Choice items made by the English teachers of MTsN-1 Model Palangka Raya as a subject research. In this study the writer will use total sample.

### **4. The Technique of Collecting Data**

#### **a. Instrumentation**

*“Instrumen merupakan alat ukur yang digunakan untuk mendapatkan informasi kuantitatif tentang variasi karakteristik variabel secara obyektif”* (Instrument is measurer used to get the qualitative information about the variation of characteristics of variable objectively).

In this research, the instrument used is summative test items of Multiple Choice made by the English teachers of MTsN-1 Palangka Raya.

### 1) *Instrument validity*

According to Tinambunan states “validity refers to extend to which the results of an evaluation procedure serve the particular uses for which they are intended”.<sup>13</sup> If the tests are valid, they can be used as the instrument for measuring the students’ ability. To measure the validity of this test, the formulation below was used:

$$r_{pbi} = \frac{Mp - Mt}{SDt} \sqrt{\frac{p}{q}}$$

Where:

$r_{pbi}$  = coefficient validity of each item

$Mp$  = mean of the right answer

$Mt$  = mean of the total score

$SDt$  = standard deviation from the total score

$p$  = proportion of the students' correct answers for each item

$q$  = proportion of the students' wrong answer for each item (1-p).<sup>14</sup>

The interpretation of correlation of coefficient is:

$r$ between	0,00 – 0,20	: no correlation
	0,20 – 0,40	: low
	0,41- 0,70	: fair
	0,71 – 0,90	: high
	0,91 – 1,00	: very high (perfect) <sup>15</sup>

<sup>13</sup>Wilmar Tinambunan, *Evaluation of Students' Achievement*, Jakarta: Departemen Pendidikan dan Kebudayaan, 1988, p. 11.

<sup>14</sup>Anas Sudijono, *Pengantar Statistik Pendidikan*, Jakarta: Raja Grafindo Persada, 2003, p. 245-246.

To determines the Mt and SDt as below:

1. Determining Mean Total (Mt) by formula:

$$Mt = \frac{\sum X_t}{N}$$

2. Determining Standard Deviation Total (SDt) by formula:

$$SDt = \sqrt{\frac{\sum X_t^2}{N} - \left(\frac{\sum X_t}{N}\right)^2}$$

There is an example of calculating item validity, it is known:

$$\begin{aligned} Mt &= \frac{\sum X_t}{N} \\ &= \frac{362}{10} \\ &= 36.2 \end{aligned}$$

$$\begin{aligned} SDt &= \sqrt{\frac{\sum X_t^2}{N} - \left(\frac{\sum X_t}{N}\right)^2} \\ &= \sqrt{\frac{14796}{10} - \left(\frac{362}{10}\right)^2} \\ &= \sqrt{1479,6 - 1310,44} \\ &= \sqrt{169,16} \\ &= 13,01 \end{aligned}$$

---

<sup>15</sup>M. Ngalim Purwanto, *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Cet. 9. Bandung: Remaja Rosdakarya, 2000, p.139.

So that the result of validity, as follows:

$$\begin{aligned}
 r_{phi} &= \frac{MP - Mt}{SDt} \sqrt{\frac{p}{q}} \\
 &= \frac{40,1 - 36,2}{13,01} \sqrt{\frac{0,7}{0,3}} \\
 &= \frac{3,9}{13,01} \times 1,5 \\
 &= 0,3 \times 1,5 \\
 &= 0.45
 \end{aligned}$$

Based on the criteria that have been established above, the coefficient correlation of  $r_{phi}$  was 0.45. Therefore, the item was fair validity.

Here are the kinds of validity are used:

#### a) Content Validity

Content validity refers to the extent to which a test measures a representative sample of the domain tasks under consideration.<sup>16</sup> The content that has been taught to the students must be based on the curriculum and syllabus. The purposed was to measure the agreement between the content of the test item and the curriculum or materials given by the teacher. The material is selected from *Bahasa Inggris untuk SMP dan MTs (PR)*, Intan Pariwara.

The following table showed the content specification of the instrument where the items distributed to the students.

---

<sup>16</sup>E.N. Gronlund, *Measurement and Evaluation in Teaching (fourth edition)*, New York: Macmillan Publishing Co., Inc. 1981, p. 90.

**Table 1**  
**Content Specification of test items**

<b>NO</b>	<b>Indicators</b>	<b>Number of items</b>	<b>Total of test items</b>
1	Asking for, giving and refusing service	1,3,18	3
2	Asking for, giving and refusing things	4	1
3	Descriptive	19	1
4	Announcements	11,15-17,23,26	5
5	Lost and found ads	-	-
6	Adjectives	27-30	4
7	Comparative and superlative adjectives	12	1
8	The simple present tense	38-40	3
9	Congratulations	2,25	2
10	Complimentings	22	1
11	Recount	8-9	2
12	Letters	20-21,32-34	5
13	The simple past tense	10,37	2
14	The past continuous tense	-	-
15	Approving and denying facts	-	-
16	Agreeing and disagreeing	6	1
17	Asking for and giving opinions	7,36	2
18	Inviting and the responses	5,13-14,35	4
19	Congratulation cards	-	-
20	relating verb	31	1
21	Indefinite pronouns	-	-
22	Adverbs of frequency	-	-
<b>Total</b>			<b>40 items</b>

#### b) Construct Validity

Construct Validity refers to the extend to which the results of a data collection process can be interpreted in terms of underlying psychological



construct. Construct validity often necessitates an extremely complicated process of validation.

Thus, the validity of a test is the extent to which the test measures what is intended to measure. So, to measure the validity the writer use the content validity and the writer used construct validity for validations of instruments.

## 2) Instrument Reliability

The reliability of test should be consistent and stable. Gronlund states that “reliability refers to how consistent test scores and other evaluation results are from one measurement to another”.<sup>17</sup> So, reliability is a necessary characteristic of measurement of any good test for it to be valid at all and a test must be reliable as a measuring instrument. The writer uses KR-20 formula to measure the reliability of the test.

To know the level of test validity and reliability can be seen on the value below:

r	between	0,00 – 0,20	: no correlation
		0, 20 – 0, 40	: low
		0, 41- 0, 70	: fair
		0, 71 – 0, 90	: high
		0, 91 – 1, 00	: very high (perfect) <sup>18</sup>

If the value is high, it can be concluded that the level of test validity and reliability are high. Meanwhile, if the value is poor, the level of test validity and reliability are poor.

---

<sup>17</sup>E.N. Gronlund, *Measurement and Evaluation in Teaching*, p. 117.

<sup>18</sup>M. Ngalim Purwanto., *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*. Bandung: PT Remaja Rosdakarya, 2006, p.139.

### 3) Discriminating Power

Discriminating power used to discriminated between students who have high and low intelligent, so a good test should have discriminating power. The writer used the formula stated by Purwanto:

$$DP = \frac{U - L}{1/2T}^{19}$$

Which:

DP = discriminating power or item discrimination

N = numbers of sample

U = number of correct answer from upper group

L = number of correct answer from lower group

T = number of upper and lower group

Criterion:

DP: 0, 00-0, 20 : low

DP: 0, 20-0, 40 : fair

DP: 0, 40-0, 70 : good

DP: 0, 70-1, 00 : Very good

### b. Documentation

This technique is used to collect written data which are related to the completeness of the data. Arikunto says that “*metode dokumentasi, yaitu mencari data mengenai hal-hal atau variable yang berupa catatan, transkrip, buku, surat*

---

<sup>19</sup>*Ibid.*, p. 120.

*kabar, majalah, prasasti, notulen rapat, legger, agenda dan sebagainya*".<sup>20</sup>

(Documentation method is looking for the data about things or variable that shaped note, transcript, book, newspaper, magazine, epigraphy, notes of meeting, leger, agenda, and so forth).

According to Sukardi "*pada teknik ini, peneliti dimungkinkan memperoleh informasi dari bermacam-macam sumber tertulis atau dokumen yang ada pada responden atau tempat, dimana responden bertempat tinggal atau melakukan kegiatan sehari-harinya*".<sup>21</sup> ( in this technique, researcher is probably to get various information from written various sources or document of respondents or place, which respondent is placed or his daily activities). This technique was used to collect written data which were related to this research.

The data would like to be obtained are the English summative test items of Multiple Choice made by the English teachers of MTsN-1 Model Palangka Raya.

Some of those documents will be needed, they are:

1. Summative test of Multiple Choice made by the English teachers
2. Syllabus and curriculum that used by the teachers
3. Answer sheet that had been answered by the students.

## **5. The Technique of Processing Data**

Processing data is "*suatu proses dalam memperoleh data ringkasan atau angka ringkasan dengan menggunakan cara-cara atau rumus-rumus tertentu*".<sup>22</sup>

(a process to find the data or summary of numeral which is used the procedures or

---

<sup>20</sup>Suharsimi Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktek*, edisi revisi VI. Jakarta: Rineka Cipta. 2006. p. 231.

<sup>21</sup>Sukardi, *Metodologi Penelitian Pendidikan: Kompetensi dan Praktiknya*, Jakarta: PT. Bumi Aksara, 2003, p. 81.

<sup>22</sup>M. Iqbal hasan, *Pokok-pokok Materi Metodologi Penelitian dan Aplikasinya*, Jakarta: Ghalia Indonesia, 2002, p. 89.

certain formulations). According to Hadeli there are seven steps to process data such as “*editing, coding, scoring, tabulation, mengolah atau menghitung data dengan statistik deskriptif, membuat interpretasi hasil pengolahan tersebut dalam bentuk pernyataan verbal, analisis data lebih lanjut untuk uji hipotesa*”.<sup>23</sup> (editing, coding, scoring, tabulation, make or count data with the descriptive statistic, to make the interpretation of result in verbal question, continuos to data analysis to exam the hypothesis).

Before the writer analyzing the data, there are some data processing procedures will be done. They are explained as bellow:

1. **Editing:** to check the completeness of the names or identities of sample, to check amount of paper of test items and the contents of samples' answer in accordance with the data needed.
2. **Coding:** to change the name of sample with code such as A1, A2, and so on.
3. **Scoring:** to give score to the students' answers based on the key answers.
4. **Tabulating:** to place the data in a certain table.
5. **Analyzing:**
  - Analyzing the validity (content validity) of the test item.
  - Analyzing the reliability of each item by using KR20 formula.
  - Analyzing the discriminating power of each item.

The total numbers of the test were 40 items.

---

<sup>23</sup>Hadeli, *Metode Penelitian Kependidikan*, Jakarta: PT. Ciputat Press, 2006, p. 91.

## 6. The Technique of Data Analysis Procedures

According to Sugiyono, data analysis is “*kegiatan setelah data dari seluruh responden atau sumber data lain terkumpul*”.<sup>24</sup> ( a process after finishing the data from respondent or another sources). Then, Bogdan in Sugiyono states that “data analysis is the process of systematically searching and arranging the interview transcripts, field notes and other materials that you accumulate to increase your own understanding of them and to enable you to present what you have discovered to others”.<sup>25</sup>

Sugiyono states that data analysis is performed by quantitative and qualitative analysis.<sup>26</sup> According to Hasan, there are two instruments to analyze data; they are statistic analysis and non statistic analysis.<sup>27</sup>

Based on the statements above, the writer used to analyze data was statistic analysis, here are some procedures were taken to analyze the data:

1. Arranging the students' score from the highest score to the lowest score.
2. Identify the highest group and a lower group separately.
3. Determine the percentage of students who answered correctly. The formula is as follow:

$$S = \frac{B}{N} \times 100$$

---

<sup>24</sup>Sugiyono, *Metode Penelitian Pendidikan: Pendekatan Kuantitatif, kualitatif, dan R&D*, Bandung: Alfabeta, 2007, p. 207.

<sup>25</sup>*Ibid.*, p. 334.

<sup>26</sup>Sugiyono, *Statistika Untuk Penelitian*, Bandung: Alfabeta, 2004, p. 14.

<sup>27</sup>M. Iqbal hasan, *Pokok-pokok Materi Metodologi Penelitian dan Aplikasinya*, Jakarta: Ghalia Indonesia, 2002, p. 98.

In which:

S = the percentage who answered the item correctly

B = the total number who answered the item correctly

N = the total number of test item<sup>28</sup>

In processing the data, the writer also will use three criteria of the good test:

#### a. Validity

The validity of a test is the extent to which the test measures what is intended to measure. If the test items are valid, so they can be used to measure the ability of the students. To measure the validity, the writer uses the content validity and constructs validity.

##### 1) *Content Validity*

The writer uses the curriculum and syllabus of English for Junior High School to measure the content validity of the test items. There are the analysis procedures of content validity:

- a) Preparing the curriculum and syllabus used by the teacher.
- b) Comparing the curriculum and syllabus material with the content of the English summative test items.
- c) Tabulating the result.

##### 2). *Construction Validity*

Construct Validity refers to the extend to which the results of a data collection process can be interpreted in terms of underlying psychological

---

<sup>28</sup>Departemen Pendidikan Nasional, *Kurikulum 2004 Sekolah Menengah Pertama (SMP)*, Jakarta: 2003, p. 26.

construct. Construct validity often necessitates an extremely complicated process of validation.

In this case, the analysis of the construction validity test has done by seeing the syllabus of English subject at MTsN-1 Model Palangkaraya which use competency based curriculum 2004. See the table (appendix 2) to know the appropriateness of the test with syllabus.

#### b. Reliability of the Test

The writer uses KR-20 formula to measure the reliability of the test:

$$r_{11} = \left( \frac{n}{n-1} \right) \left( \frac{St^2 - \sum p_i q_i}{St^2} \right)$$

Where:

$r_{11}$  : Coefficient reliability of test

$n$  : Number of test items.

$1$  : Constant multiplier.

$St^2$  : Total of Varian

$p_i$  : Mean of the correct answer.

$q_i$  : Mean of the wrong answer.

$\sum p_i q_i$  : Total multiplication between  $p_i$  and  $q_i$ .<sup>29</sup>

---

<sup>29</sup>Anas Sudijono, *Pengantar Evaluasi Pendidikan*, Jakarta: PT. Raja Grafindo Persada, 2005, p. 252-253.

The following is the example calculation to find out the test reliability. It is known:

$$X = 634$$

$$\Sigma X^2 = 41456$$

$$\begin{aligned}\Sigma Xt^2 &= \Sigma X^2 - \left( \frac{(X)^2}{N} \right) \\ &= 41456 - \left( \frac{634^2}{10} \right) \\ &= 41456 - 40195.6 \\ &= 1260.4\end{aligned}$$

$$\begin{aligned}st^2 &= \left( \frac{\Sigma Xt^2}{N} \right) \\ &= \frac{1260.4}{10} \\ &= 126.04\end{aligned}$$

Now, it is known:  $n = 10$ ,  $\Sigma p_i q_i = 4.403$ ,  $St^2 = 126.04$

$$\begin{aligned}r_{11} &= \left( \frac{10}{10-1} \right) \left( \frac{126.04 - 23.2}{126.04} \right) \\ &= \left( \frac{10}{9} \right) \left( \frac{102.84}{126.04} \right) \\ &= 1.11 \times 0.82 \\ &= \mathbf{0.91}\end{aligned}$$



To know the level of reliability can be seen on the value below:

- r between    0,00 – 0,20    : no correlation  
                   0, 20 – 0, 40    : low  
                   0, 41- 0, 70    : fair  
                   0, 71 – 0, 90    : high  
                   0, 91 – 1, 00    : very high (perfect)<sup>30</sup>

So, it is found that  $r_{11} = 0.90 > 0.70$ , it is interpreted very high reliable.

### c. Discriminating Power

Discriminating power is the ability of a test item to discriminate between high intelligence students and low intelligence students.<sup>31</sup> There are some steps that the writer did in analyzing whether the test had discriminating power or not:

1. Preparing the data
2. Tabulating the data from the highest to the lowest score
3. Calculating the discriminating power of each item by using following formula:

$$DP = \frac{U - L}{1/2T}$$

Which:

DP = discriminating power or item discrimination

N = numbers of sample

U = number of correct answer from upper group

L = number of correct answer from lower group

T = number of upper and lower group.<sup>32</sup>

---

<sup>30</sup>M. Ngalim Purwanto, *Teknik Evaluasi* p.139.

<sup>31</sup>*Ibid.*, p. 26.

Criterion:

DP: 0, 00-0, 20 : low

DP: 0, 20-0, 40 : fair

DP: 0, 40-0, 70 : good

DP: 0, 70-1, 00 : Very good

## **I. Framework of the Discussion**

This thesis design consists of four chapters. Chapter 1 consists of Background of the Study, Problems of the Study, Objectives of the Study, Significances of the Study, Definition of the Concept, Delimitation of the Study, Variable of the Study, Research Methodology (Method of the Study, The subject of the Study, The Technique of Collecting Data, The Technique of processing Data, The Data Analysis Procedures), and Framework of the Discussion.

Chapter 2 consists of Test, Uses and functions of test, Kinds of test, the principles in writing Multiple Choice, the strengths and the weaknesses, The characteristics of good test, test procedure, scoring test and the review of research finding.

Chapter 3 consists of the quality of English Summative test of Multiple Choice made by the English teachers of MTsN-1 Palangka Raya. Measuring the validity, reliability and the discriminating power of the test items.

Chapter 4 consists of Conclusion and Suggestion.

---

<sup>32</sup>M. Ngalim Purwanto, *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*. Bandung: PT Remaja Rosdakarya, 2006, p. 120.

# **CHAPTER II**

## CHAPTER II

### REVIEW OF LITERATURE

#### A. Test

Test represents one particular measurement technique. It is an instrument or systematic procedure for measuring a sample of behavior. There are definitions of test based on some experts. Indrakusuma in Arikunto says "*tes adalah suatu alat atau prosedur yang sistematis dan objective untuk memperoleh data-data atau keterangan-keterangan yang diinginkan tentang seseorang, dengan cara yang boleh dikatakan cepat dan tepat*".<sup>19</sup> (test is an instrument or objective and systematically procedures to obtain the data or explanation about someone, by the way that can be said fast and accurate). According to Thoha "*tes adalah alat pengukuran berupa pertanyaan, perintah, dan petunjuk yang ditujukan kepada testee untuk mendapatkan respon sesuai dengan petunjuk itu*".<sup>20</sup> (test is an instrument includes question, instruction, and direction pointed to testee to get respond according to the direction). The other opinion is from Tinambunan who defines that "a test is a set of questions, each of which has a correct answer, that examinees usually answer orally or in writing".<sup>21</sup> And a test will be given by the teacher to measure the student's achievement.

Based on the statements above, it can be concluded that the test is an instrument or systematic procedure in measuring the achievement of student during the teaching

---

<sup>19</sup>Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, Jakarta: Bumi Aksara, 2003, p. 32.

<sup>20</sup>M.C. Thoha, *Teknik Evaluasi Pendidikan*, Jakarta: Raya Grafindo Persada, 2003, p. 43.

<sup>21</sup>Wilmar Tinambunan, *Evaluation of Students' Achievement*, Jakarta: Departemen Pendidikan dan Kebudayaan, 1988, p. 6.

learning process. In this study the test is used in education area or school. So, the result of the test is benefit not only for the students but also for improving and progressing of the school.

## **B. Uses and Functions of Test**

### **1. The uses of Test**

The significance of evaluation in educational field is there is possibility for the evaluator to get the information about the result achieved in keeping the implementation of educational programmed.

According to Sudijono, there are two uses of test:

- a. *Terbukanya kemungkinan untuk dapat diketahuinya relevansi antara program pendidikan yang telah dirumuskan, dengan tujuan yang hendak di capai.* (there is possibility to know the relevance between the educational programmed that had been formulated with the goal achieved).
- b. *Terbukanya kemungkinan untuk dapat dilakukannya suatu perbaikan, penyesuaian dan penyempurnaan program pendidikan yang dipandang lebih berdaya guna dan berhasil guna. Sehingga tujuan yang hendak dicita-citakan dapat dicapai dengan hasil yang sebaik-baiknya.* (there is possibility to do improvement, adaptation and finishing the educational programmed that seen more effective and efficiency so the goal that wanted can be achieved with the best result)<sup>22</sup>

Based on the explanations above, the uses and the functions of the test are to get information that show students' progress after following the teaching learning process and for the teacher they will know the weaknesses and the difficulties and the effectiveness of teaching learning process.

### **2. The Functions of Test**

The Function of test in education cannot be discharged from the purpose of evaluation itself. The purpose of evaluation in education is to get the verification data

---

<sup>22</sup> Anas Sudijono, *Pengantar Evaluasi Pendidikan*, Jakarta: Raja Grafindo, 2007, p. 17.

that will showed until where the level of ability and student's success in attainment of curricular. Beside that, it also can be used by all teachers and education supervisor to measure or assess until where the effectiveness of teaching experience, learning activities and teaching methods that used.<sup>23</sup>

According to Sudijono there are two functions of test:

1. *Sebagai alat ukur peserta didik. Dalam hubungan ini tes berfungsi untuk mengukur perkembangan atau kemajuan yang telah dicapai oleh peserta didik setelah mereka menempuh proses belajar mengajar dalam jangka waktu tertentu.* (as a means of measuring educative participant. In this relation, a test has function to measure the progress reached by educative participant after the educative participants learned in selected time).
2. *Sebagai alat pengukur keberhasilan program pengajaran, sebab dengan melalui tes tersebut akan dapat diketahui sudah seberapa jauh program pengajaran yang telah ditentukan, telah dapat dicapai.* (as a means to measures learning program, because through the test, it can be known how far instruction program have been achieved).<sup>24</sup>

According to Heaton the functions of test are:

1. to measure the students performance in classroom.
2. to diagnose the students weaknesses and difficulties.
3. to evaluate the effectiveness of the syllabus as well as the methods and material used.
4. to provide the students with an opportunity to show their ability to recognize and produce correct form of the language.<sup>25</sup>

---

<sup>23</sup>Rini Yulia, "Kemampuan Guru Menentukan Nilai Akhir Mata Pelajaran PAI di SMA Palangka Raya", *skripsi*, Palangka Raya: STAIN, 2006, p. 17.

<sup>24</sup>Anas Sudijono, *Pengantar Evaluasi Pendidikan*, Jakarta: Raja Grafindo, 2007, p. 67.

<sup>25</sup>J.B Heaton, *Writing English Language Test*, London: Longman Group Ltd, 1979, p. 2.

### C. Kinds of Test

The followings are the kinds of test, they are:

#### 1. Based on Organization

##### a. *Standardize Test*

Gronlund states that “Standard test measure the common objectives of a wide variety of schools have standard procedures of administration and scoring and provide norms for interpreting the scores”.<sup>26</sup>

##### b. *Teacher Made Test*

Slameto says “*tes buatan guru yaitu tes yang dibuat oleh guru untuk keperluan penilaian guru tersebut terhadap siswanya. Tes ini terutama tes hasil belajar dan biasanya berlaku untuk satu sekolah, bahkan kadang-kadang hanya untuk satu kelas saja*”.<sup>27</sup> (test made by the teacher is a test that made by the teacher for assessing the student. This test especially in test result of learning and usually valid for one school, even sometimes for only one class).

#### 2. Based on Method of Teaching

##### a. *Norm-Referenced Test*

Gronlund states “norm-referenced test (NRTs) is design to rank pupil in order of achievement, from high to low, so that decision based on the relative achievement (e.g. selection, grouping, grading) can be made with greater confidence”.<sup>28</sup> It means NRT are made to compare test takers to each other. On an NRT driving test, test-takers would be

---

<sup>26</sup>E.N. Gronlund, *Measurement and Evaluation in Teaching (fourth edition)*, New York: Macmillan Publishing Co., Inc. 1981, p. 322.

<sup>27</sup>Slameto, *Evaluasi Pendidikan*. Cet.2. Jakarta: Sinar Grafika Offset, 2001, p. 32.

<sup>28</sup>E.N. Gronlund, *Measurement and Evaluation in Teaching (fourth edition)*, 1981, p. 20.

compared as to who knew most or least about driving rules or who drove better or worse. Score would be reported as a percentage rank with half scoring above and half below the mid-point.

#### **b. Criterion Reference Test**

According to Purwanto, states that “*criterion-referenced test (CRTs) adalah sekumpulan soal atau items yang secara langsung mengukur tingkah laku- tingkah laku yang dinyatakan didalam seperangkat tujuan behavioral atau performance objectives*”.<sup>29</sup> (criterion-referenced test is a group of items which directly measure behaviors that expressed in a set the target of behavioral or of performance objective).

In education, CRTs usually are made to determine whether a student has learned the material taught in a specific grade or course. On a standardized CRT (one taken by students in many school), the passing or “cut-off” score is usually set by a committee of experts, while in a classroom the teacher sets the passing score.

Based on the explanation above, it can be concluded that norm-referenced test is designed to rank pupil from their achievement and criterion reference test is to describe the students’ performance in terms of specified domain of instructionally relevant tasks. If standardized tests are used at all, CRTs make more sense for schools than do NRTs. However, they should be based on relevant, high quality standards and curriculum and should make the least possible use of multiple choice and short answer question. As with all tests, CRTs and NRTs, no matter what they are called, should not control curriculum and instruction, and important decisions about students, teachers or schools should not be based solely or automatically on test score.

---

<sup>29</sup>M. Ngalim Purwanto, *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Cet. 9. Bandung: Remaja Rosdakarya, 2000, p. 27.



### 3. Based on Its Function

There are some functions of test based on Gronlund, they are:

- a. **Placement Test** is a test to measure the students' basic ability and to place them based on their ability in a right level or suitable group.
- b. **Formative Test** determines learning progress, provide feedback for reinforcement of learning and correct learning errors.
- c. **Diagnostic Test** is designed to diagnose the students' difficulties and weaknesses during the teaching learning process.
- d. **Summative Test** is given at the end of a course or unit of instruction and the results are primarily for assigning grades or for certifying pupil mastery of the instructional objectives.<sup>30</sup> Another definition is based on Cole in Halimi "summative evaluation is typically conducted at the end of a course or unit of instructional". The teacher and the students should participate in the final evaluation process.<sup>31</sup> It can be concluded that summative evaluation is a method of judging the worth of a program at the end of the program activities. In this study the summative evaluation is focused on the out come of the student. So, the process of teaching learning is demanded to reach the goal of this evaluation. Whether or not the goal of this evaluation reached. Can be seen after doing the summative evaluation.

### 4. Based on the Form of the Test

There are two kinds of test based on the form of the test, they are:

---

<sup>30</sup>E.N. Gronlund, *Measurement and Evaluation in Teaching (fourth edition)*, New York: Macmillan Publishing Co., Inc. 1981, p. 21.

<sup>31</sup>Akhmad Halimi, *Summative Evaluation Problems of English Lesson Faced by The Second Years Students of MTs-Muslimat NU Palangka Raya*, skripsi, 2008, td.

### **a. Written Test**

Thoha says that “*tes tertulis adalah tes yang soal dan jawaban yang diberikan kepada siswa berupa bahasa tulisan*”.<sup>32</sup> (written test is a test which is the items and the answer that have been given to the student is in the form of written language). It means the testers have to answer the test by writing the answers on the paper and it is called the students’ answer sheet.

### **b. Oral Test**

Oral test is a test where the questions are from the tester/teacher orally and the testers/students have to answer orally too. Based on Heaton, “oral test is testing ability to speak”.<sup>33</sup> This test is usually in the form of speaking.

From the statement above, oral test is a test where the students accept the question from the teacher orally and the students should answer orally too.

### **c. Performance test**

Performance test is an act of the student’s performs a play, drama or some other entertainment.

## **5. Based on the Scoring System**

### **a. Subjective Test**

A subjective test is a test where the scores are influenced by the opinion or judgment of the person doing the scoring. According to Sudijono, “*tes subyektif adalah salah satu jenis tes hasil belajar yang memiliki karakteristik sebagaimana dikemukakan berikut*”: (subjective test is one of the test kinds for the studying result that has characteristics as below:)

---

<sup>32</sup>M.C. Thoha, *Tehnik Evaluasi Pendidikan*, Jakarta: Raya Grafindo Persada, 2003, p. 54.

<sup>33</sup>J.B. Heaton, *Writing English Language Test*, London: New Jersey, 1974, p. 83.

1. *Tes tersebut berbentuk pertanyaan atau perintah yang menghendaki jawaban berupa uraian atau paparan kalimat pada umumnya cukup panjang.* (the test formed question or command which need answer such as essay or explanation of sentence, generally is long enough).
2. *Bentuk-bentuk pertanyaan atau perintah itu menuntut kepada testee untuk memberikan penjelasan komentar, penafsiran, membandingkan, membedakan dan sebagainya.* (the form of question or command demands to the testee for giving explanation, comment, interpretation, comparison, and differentiation and so on).
3. *Jumlah butir soalnya umumnya terbatas yaitu berkisar antara lima sampai dengan sepuluh butir.* (the total of the tests are limited about five to ten numbers).
4. *Pada umumnya butir-butir soal tes uraian itu diawali dengan kata "Jelaskan...", "Terangkan...", Uraikan...", "Mengapa...", "Bagaimana..." atau kata-kata lain yang serupa dengan itu.* (generally, the number of the item test is begin with the words : "Explanation...", "Clarify...", "Mention...", "Why...", "How...", or the other words which are similar).<sup>34</sup>

Examples:

- 1) Why are vitamins and minerals needed?
- 2) What do you think of the poem?

#### **b. Objective Test**

According to Tinambunan says that, "Objective test is a test which is highly structured and requires the pupils to supply a word or two or to select the correct answer

---

<sup>34</sup>Anas Sudijono, *Pengantar Evaluasi Pendidikan*, Jakarta: PT. Raja Grafindo Persada, 2005, p. 100.

from among a limited number of alternatives”.<sup>35</sup> Another definition is from Thoha, “*tes obyektif adalah tes yang itemnya dapat dijawab dengan memilih jawaban yang sudah tersedia sehingga peserta menampilkan keseragaman data, baik bagi yang menjawab benar maupun mereka yang menjawab salah*”.<sup>36</sup> (objective test is a test which its item answerable by choosing the available answer. So, that participant presents the similarity of the data both of student who has true answer or the student who has false answer).

From the explanation above, it can be concluded that the objective test is one kind of tests formed by Multiple Choice and the students can choose the right answer to answer the item or question.

Objective test items can be classified into supply type and selection type.

#### **1). Supply type.**

The supply type test requires pupil to supply the answer, this type is also known as ‘short answer’ or ‘completion’. According to Gronlund states, “supply test is some test requires examinees to supply the answer. (ex. Essay test)”.<sup>37</sup> The short answer and the completion item are both supply test.

Examples:

1. I ..... a new car, yesterday. (bought).
2. I’m ..... a song. (singing).
3. I ..... to school everyday (go).

---

<sup>35</sup>Wilmar Tinambunan, *Evaluation of Students’ Achievement*, Jakarta: Departemen Pendidikan dan Kebudayaan, 1988, p. 55.

<sup>36</sup>M. C. Thoha, *Teknik Evaluasi*, p. 66.

<sup>37</sup>E.N. Gronlund, *Measurement and Evaluation in Teaching (fourth edition)*, New York: Macmillan Publishing Co., Inc. 1981, p. 144.

## 2). *Selection type*

Gronlund state, “selection test require them to select the correct response from a set of alternatives”. (ex. Multiple Choice test).<sup>38</sup>

This type can be further subdivided into :

### a). *Matching*

Matching exercises typically consist of two columns, each item in the first column to be paired with an alternative in the second column.

Examples:

Match these words with their meaning!

- |           |           |
|-----------|-----------|
| 1. Ugly   | a. Pintar |
| 2. Clever | b. Malas  |
| 3. Lazy   | c. Jelek  |

### b). *Multiple Choice*

A multiple choice test is made up of items which present two or more responses, only one of them is correct or definitely better than the others. Multiple choice test consist of two part, they are “stem” is the introductory question or incomplete statement at the beginning of each item and this is followed by the options, then the “option” which consist of the answer -- the correct option-- and “destructors”—the incorrect but (it hopes) tempting options.<sup>39</sup> From the explanation above, it can be concluded that the objective test is one kind of test formed by Multiple Choice and the students can choose the right answer to answer the item or question.

---

<sup>38</sup>E.N. Gronlund, *Measurement and Evaluation in Teaching (fourth edition)*, New York: Macmillan Publishing Co., Inc. 1981, p. 18.

<sup>39</sup> [www.park\\_edu.com](http://www.park_edu.com)



1. Make the instructions for each type of question simple and brief.
2. Use simple and clear language in the questions. If the language is difficult, students who understand the material but who do not have strong language skills may find it difficult to demonstrate their knowledge. If the language is ambiguous, even a student with strong language skills may answer incorrectly if his or her interpretation of the question differs from the instructor's intended meaning.
3. Write items that require specific understanding or ability developed in that course, not just general intelligence or test wise ness.
4. Do not suggest the answer to one question in the body of another question. This makes the test less useful, as the test wise student will have an advantage over the student who has an equal grasp of the material, but who has less skill at taking tests.
5. Do not write questions in the negative. If you must use negatives, highlight them, as they may mislead students into answering incorrectly.
6. Specify the units and precision of answers. For example, will you accept numerical answers that are rounded to the nearest integer.

The most commonly used type of question is the Multiple Choice question. Multiple-choice questions are more easily and objectively graded than essay questions and are more difficult to answer correctly without the required knowledge than true-false questions.

Zaini in Dina, states some principles in writing multiple choice items, here are the general principles for writing multiple choice items:

1. The main problem should be placed on stem.
2. Avoid repeating words from the stem and in the correct response option.
3. Use clear, precise and simple language.

4. When utilizing an incomplete statement, place blanks at the end rather than the beginning or middle of stem.
5. .Make all response options equivalent in length, style, and structure.
6. Avoid using scientific terminology.
7. Make all response options is homogeny.
8. Avoid make the correct answer is longer than wrong answer.
9. Avoid the indicators for the correct answer.
10. Avoid “all of the above” or “none of the above” response options as students may correctly answer the question.
11. Include three to five response options in each multiple choice items.
12. Avoid using absolute terminology (always, never, etc).
13. Try to restrict the use of negatives in the stem. When negatives are necessary highlight them (bold, underline, capital letters) to draw the student attention to the shift in focus.<sup>40</sup>

Multiple Choice questions, however, are probably the most difficult type of question to construct.

#### **D. The Strengths and The Weaknesses**

As we know, generally the tests made by the teachers are usually divided into two general types based on the scoring system. They are subjective and objective test, every test has the strengths and the weaknesses. The strengths and the weaknesses of each test, are:

---

<sup>40</sup>Dina Hadma., “The Validity and Reliability of Multiple Choice Test in Measuring Student’s Achievement at MAN Model of Palangka Raya”, *Skripsi*, Palangka Raya: STAIN, 2008, p. 19, t.d.



## 1. Subjective Test

### a. *The Strengths*

- 1) The essay item is the most effective in assessing complex learning outcomes because the students are given a freedom to decide how to approach the problem, what factual information to use, how to organize the answers, and what degree of emphasis to give to each aspect of the response. Thus, the essay question is especially useful for measuring the ability of organizing, integrating and expressing idea.
- 2) Constructing essay questions are relatively easy.
- 3) The possibility of guessing is minimized.
- 4) Constructing essay questions require less time.<sup>41</sup>

### b. *The Weaknesses*

- 1) Scoring essay question is difficult.
- 2) The unreliability of scoring.
- 3) Limited sample of total instructional content.
- 4) Judgment of student responses requires much time.
- 5) Bluffing.<sup>42</sup>

---

<sup>41</sup>Wilmar Tinambunan, *Evaluation of Students' Achievement*, Jakarta: Departemen Pendidikan dan Kebudayaan, 1988, p. 87.

<sup>42</sup>*Ibid.*, p. 89.

## 2. Objective Test

### a. Supply Type (Completion Items)

#### 1) *The Strengths*

- a) The test can be used effectively to measure the recall of terms, dates, names, and generalizations.
- b) Can be used at almost all grade levels, but it seems to be especially appropriate at the intermediate level.
- c) Its importance as a testing technique decreases.
- d) It presents to the examinee a problem to be solved by the recall and recording of an appropriate word, or phrase.
- e) The item should be concise, unambiguous, and grammatically correct.<sup>43</sup>

#### 2) *The Weaknesses*

- a) The student might respond with any one of a number of answers that could be defended as appropriate.
- b) The problem of spelling must be also considered carefully.<sup>44</sup>

### b. Selection Item

#### 1) Matching Item.

##### a) The Strengths

- The matching item has been widely used in workbooks and tests accompanying textbooks.
- The effect of guessing is correctly.<sup>45</sup>

---

<sup>43</sup>*Ibid.*, p. 61.

<sup>44</sup>*Ibid.*, p. 62.

b) The Weaknesses

- The pupil is not free to redefine the problem or to organize and present the answer in his own words.
- If too many clues are given in the attempt to establish a single frame of reference, the item will be too easy.
- The item will be ambiguous and many levels itself to several correct responses, if an insufficient number of clues are presented.<sup>46</sup>

2) Multiple Choice.

a) The Strengths

- The Multiple Choice can be more effectively to measure many of the simple learning outcomes measured by the supply test, the true-false item, and the matching item.
- It can measure a variety of the more complex learning outcomes in the knowledge, understanding and application areas.
- Flexibility.
- In the Multiple Choice the higher quality items usually found.
- The scoring is quick, easy and accurate.<sup>47</sup>

b) The Weaknesses

- *Pokok soal tidak cukup jelas sehingga terdapat kemungkinan ada lebih dari satu jawaban yang benar* (the content of items is not clear enough. Therefore, there is possibility that there is more than one correct answer).

---

<sup>45</sup>*Ibid.*, p. 65.

<sup>46</sup>*Ibid.*

<sup>47</sup>*Ibid.*, p. 56.

- *Kadang-kadang jawaban soal dapat diketahui siswa meskipun belum diajarkan karena adanya petunjuk jawaban yang benar, atau karena butir soal itu mengukur sikap dan bukan mengukur pengetahuan. (sometimes the answer of the items can be known by the student although it is not been taught yet. It is caused by guide of correct answer, or because the item is to measure attitude and is not to measure knowledge).*
- *Sampai suatu tingkat tertentu keberhasilan atas suatu jawaban dapat diperoleh melalui tebakan. (the answer can be found by blind guessing the options).*
- *Sulit membuat pengecoh (distractor) yang berfungsi, yakni yang mempunyai peluang cukup besar untuk dipilih oleh siswa. (it is difficult to make distractors that have function with big chance to be selected by student).*
- *Membutuhkan waktu yang lama untuk menulis soal-soalnya, tiap soal terdiri atas pokok soal dan pilihan jawaban. Membuat soal dan terutama membuat pengecoh memerlukan pertimbangan tertentu yang membutuhkan waktu yang lama. (it needs long time to write the items. It is because each item consists of the content of item and the options).*
- *Siswa cenderung mengembangkan cara belajar terpisah-pisah menurut bunyi tiap soal. Hal ini lebih jelas terlihat apabila siswa belajar dari butir-butir yang sudah ada, siswa tidak lagi membuat suatu keseluruhan konsep tentang bahan yang dipelajarinya. (students tend to develop their learning way separately according to sound of each item. It can be seen clearer if the*

students learn from the items existing; the students do not make an overall of concept concerning materials they have learned).<sup>48</sup>

### 3) True-false

#### a) The Strengths

- *Pembuatannya mudah.* (easy to make)
- *Dapat dipergunakan berulang kali.* (it can be used repeatedly)
- *Dapat mencakup bahan pelajaran yang luas.* (it can cover wide of Lesson materials)
- *Tidak terlalu banyak memakan lembaran kertas.* (it doesn't consume many paper sheets)
- *Bagi testee, cara mengerjakannya mudah.* (for tester, the way of doing the test is easy)
- *Bagi tester, cara mengoreksinya yang mudah.* (for tester, the way of correcting the test is easy).<sup>49</sup>

#### b) The Weaknesses

- *Membuka peluang bagi testee untuk berspekulasi dalam memberikan jawaban.* (it opens the chance for the tester to take a gamble in giving their answer)
- *Sifatnya amat terbatas (hapalan).* (its character is limited)
- *Pada umumnya test obyektif jenis ini reliabilitasnya rendah, kecuali apabila butir-butir soalnya dibuat dalam jumlah yang banyak sekali.* (in general, the

---

<sup>48</sup>Suke Silverius, *Evaluasi Hasil Belajar dan Umpan Balik*, Jakarta: PT Grafindo, 1991, p. 68.

<sup>49</sup>Anas Sudijono, *Pengantar Evaluasi Pendidikan*, Jakarta: Raja Grafindo, 2005, p. 108.

objective test type for its reliability is low, except if the items made in many numbers)

- *Dapat terjadi bahwa butir-butir soal tes obyektif jenis ini dapat dijawab dengan 2 kemungkinan saja yaitu benar/salah. (it can happen that each items of the objective test type can be answered by two possibility that is correct / wrong).*<sup>50</sup>

### E. The Characteristics of Good Test

There are many statements about criteria of a good test based on some experts.

Gronlund states that “characteristics can be classified under the headings of validity, reliability and usability”.<sup>51</sup> Based on Joni in Melya states that:

*“Adapun ciri-ciri yang harus dimiliki oleh suatu test agar bisa menunaikan fungsinya sebagaimana diharapkan validita, reliabilita, tingkat kesukaran, kemampuan diskriminasi dan syarat-syarat tambahan lainnya seperti keluasan ruang lingkup, kemudahan administrasi serta scoring dan lain-lain”.*<sup>52</sup> (the characteristics which have to be had by a test is it can obtain its function likes validity, reliability, level of difficulty, discriminating power, and other additional prerequisite, such as easy of administration and also scoring and others.)

Based on the statements above, the writer only takes some criteria of a good test that are very important (validity, reliability, and discriminating power).

#### 1. Validity

The most important variable in judging adequacy of a measurement instrument is its validity. According to Tinambunan states “validity refers to extend to which the

---

<sup>50</sup>*Ibid.*, p. 109.

<sup>51</sup>E.N. Gronlund, *Measurement and Evaluation in Teaching (fourth edition)*, New York: Macmillan Publishing Co., Inc. 1981, p. 65.

<sup>52</sup>Melya Oktarini Y.B., “The Quality of English Summative Test Made by The English Teacher of SMP Kristen Palangka Raya”, *Skripsi*, Palangka Raya: Universitas Palangka Raya, 2007, p. 12, t.d.

results of an evaluation procedure serve the particular uses for which they are intended".<sup>53</sup> Thus, the validity of a test is the extent to which the test measures what is intended to measure.

There are four types of validity are commonly identified, respectively, as content validity, criterion-related validity, construct validity, and face validity. Each can be defined as the degree to which a measuring instrument accomplishes the aim associated with that type.

#### **a. Content Validity**

Content validity refers to the extent to which a test measures a representative sample of the domain tasks under consideration.<sup>54</sup> The content that has been taught to the students must be based on the curriculum and syllabus.

#### **b. Criterion-Related Validity**

Criterion-Related Validity refers to how closely performance on a data collection process is related to some others measure of performance. There are two types of criterion-related validity:

- *Predictive validity* refers to how well a data collection process predicts some future performance. Procedure is by comparing test scores with another measure of performance obtained at later date.
- *Concurrent validity* refers to how well data collection process correlates with some current criterion-usually another test. It "predicts" the present. According Purwanto defines concurrent validity as "*jika hasil suatu tes mempunyai korelasi yang tinggi*

---

<sup>53</sup>Wilmar Tinambunan, *Evaluation of Students' Achievement*, Jakarta: Departemen Pendidikan dan Kebudayaan, 1988, p. 11.

<sup>54</sup>E.N. Gronlund, *Measurement and Evaluation in Teaching (fourth edition)*, New York: Macmillan Publishing Co., Inc. 1981, p. 90.

*dengan hasil suatu alat ukur lain terhadap bidang yang sama pada waktu yang sama pula, maka dikatakan tes itu memiliki concurrent validity*".<sup>55</sup> (if the result of test has high correlation to some other valued measure performance at same current performance too. It is said the test has concurrent validity).

### **c. Construct Validity**

Construct Validity refers to the extend to which the results of a data collection process can be interpreted in terms of underlying psychological construct. Construct validity often necessitates an extremely complicated process of validation. So, a test is said has construction validity if the item is arranged to be able to measure every aspect of thinking that is written in curriculum.

### **d. Face Validity**

In face validity, you look at the operationalization and see whether "on its face" it seems like a good translation of the construct. This is probably the weakest way to try to demonstrate construct validity. It would be clearly weak evidence because it is essentially a subjective judgment call. (Note that just because it is weak evidence doesn't mean that it is wrong. We need to rely on our subjective judgment throughout the research process. It's just that this form of judgment won't be very convincing to others). We can improve the quality of face validity assessment considerably by making it more systematic. For instance, if you are trying to assess the face validity of a math ability measure, it would be more convincing if you sent the test to a carefully selected sample of experts on math ability testing and they all reported back with the judgment that your measure appears to be a good measure of math ability.

---

<sup>55</sup>M. Ngalim Purwanto., *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Cet. 9. Bandung: Remaja Rosdakarya, 2000, p. 56.



## 2. Reliability

Reliability is the second characteristics of the good test. In this study, the reliability is meant the stability of the test score. Gronlund states that “reliability refers to how consistent test scores and other evaluation results are from one measurement to another”.<sup>56</sup> So, reliability is a necessary characteristic of measurement of any good test for it to be valid at all and a test must be reliable as a measuring instrument. In interpreting and using reliability information, it is important to remember that reliability estimates refer to the results of measurement. The reliability can be seen based on some methods: test-retest, equivalent-form, split-half, and Kuder-Richardson method.

## 3. Discriminating Power

Discriminating power *adalah bagaimana kemampuan soal itu untuk membedakan siswa-siswa yang termasuk kelompok pandai (upper group) dengan siswa-siswa yang termasuk kelompok kurang pandai (lower group)*.<sup>57</sup> (is how the ability of the item to divide the students who belong to clever group (upper group) with students who belong to less clever group (lower group). So, discriminating power is the ability of a test item to discriminate between high intelligence students and low intelligence students. It is important to find out discriminating power of the item because when the teacher constructs the items, it is expected that the items are able to give the test result which shows discrimination of students' intelligence. Discriminating power of a test item can be known from the discriminating index. Discriminating index symbolized by “D” from words discriminating power. The same as difficulty index, discriminating index also

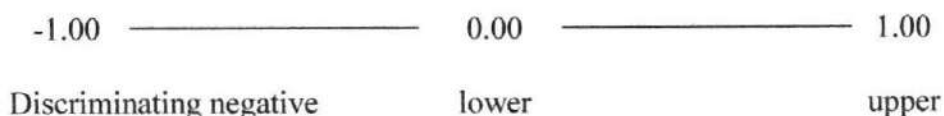
---

<sup>56</sup>*Ibid.*, p. 117.

<sup>57</sup>M. Ngalim Purwanto., *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Cet. 9. Bandung: Remaja Rosdakarya, 2000, p. 27.

range between 00.0 to 1.00. In discriminating index, there is negative sign (-). It uses the quality of tester. In this case, the negative sign (-) will be used if the upper students called lower students and lower students called upper students.

The three points of discriminating power can be drawn as follows:



So the tester can be classified into two groups, upper and lower group. If the upper group can answer the test well and the lower group has the wrong answer, the item will have D 1.00. While the upper has the wrong answer and the lower group can answer well, the item has D – 1.00, but if both groups can or cannot answer the question, the item has D 0.00 because there is discriminating power.

The advantages of discriminating power of items are:

- 1) *Untuk meningkatkan mutu setiap butir soal melalui data empiriknya. Berdasarkan indeks daya pembeda, setiap butir soal dapat diketahui apakah butir soal itu baik, direvisi, atau ditolak.* (to increase the quality of each item through its empirical data. According to index of discriminating power, each item can be known whether the item is good, need revision or rejected)
- 2) *Untuk mengetahui seberapa jauh setiap butir soal dapat mendeteksi/membedakan kemampuan siswa, yaitu siswa yang telah memahami atau belum memahami materi yang diajarkan guru.* (to know how far each item can detect/discriminate the students' ability,

it means whether the students have understood or have not understood yet about the material given by the teacher).

If an item can not discriminate the both ability of the students so the item can be suspected “possibility” like these:

- a. *Kunci jawaban butir soal itu tidak tepat.* (the key answer of the item is not appropriate).
- b. *Butir soal memiliki 2 atau lebih kunci jawaban yang benar.* (the item has two or more of key answer that true).
- c. *Kompetensi yang diukur tidak jelas.* (the competence measured is not clear).
- d. *Pengecoh tidak berfungsi.* (the distractors has not function).
- e. *Materi yang ditanyakan terlalu sulit, sehingga banyak siswa yang menebak.* (the material asked is so difficult so many students just guess the answer).
- f. *Sebagian besar siswa yang memahami materi yang ditanyakan berpikir ada yang salah informasi dalam butir soalnya.* (most of the students who understand the material asked think that there is a wrong information in the item).

## **F. Test Procedure**

There are three kinds of form of the test that usually used by the teachers in MTsN-1 Palangka Raya, they are: Written Test, Oral Test and Performance test. And, the teachers of MTsN-1 Model Palangka Raya for Summative test in academic 2008/2009 were used written test. The total numbers of the test were 40 items.

## **G. Scoring Test**

As we know, evaluation is a process of gathering the data to decide how far, in what matter and how the purpose of education has reached. In order to get subjective

decision, so it is needed certain criteria. The function of the criteria is as measurement, whether someone has fulfilled the requirement to be include as a success, clever student, up to the high class, graduated or not. The criterion is called “Scoring Standard”.

There are two kinds of Scoring Standard, they are:

1. *Standar Penilaian yang Mutlak*. (absolute scoring standard).

2. *Standar Penilaian yang Relatif*. (relatif scoring standard).

a) *Standar Penilaian yang Mutlak*. (absolute scoring standard).

This criterion is known as “*Penilaian Acuan Patokan*” in Indonesia. This standard has constant characteristic. It means that the success of students is not influenced by achievement or the group of the students.

b) *Standar yang Relatif*. (relative scoring standard).

This criterion is known as “*Penilaian Acuan Normal*” in Indonesia. Differ from Absolute Scoring Standard, in this standard the success of the students is determined by their position among the other group students who follow the evaluation.

## **H. The Review of Research Finding**

The writer also found that there is one previous study, which almost similar to her study that is The Quality of English Summative test Made by The English Teacher of SMP Kristen Palangka Raya by Melya Oktarini Y.B. 2007. She was to analyze five characteristics of a good test (validity, reliability, discriminating power, level of difficulty, and practicability).

The result of Melya’s study showed that the content validity of the English Summative test made by the English Teacher of SMP Negeri 8 for the second grade student the second semester in Academic year 2005/2006 valid. The reliability of test was

so it was not reliable. The result of discriminating power found by the writer only (8%), so it was not discriminable. Level of difficulty the result found by the writer that (44%), so the test was suitable to the level of the student, while the practicability of the test, the writer found that there were 12 items (80%) of question suitable to the indicator of the practicability, so it belonged to the practicable category.

# **CHAPTER III**

### CHAPTER III

#### RESULT OF THE STUDY

##### **A. The English Summative Test in MTsN-1 Model Palangka Raya.**

The English summative test made by the English teacher of MTsN-1 Model Palangka Raya to the eighth grade student at the first semester in academic year 2008/2009 was made by two teachers. There were 40 (forty) items and all of items presented in Multiple Choice form. Each item had four alternative options (a, b, c, and d). The teachers used the book entitled *Bahasa Inggris untuk SMP dan MTs kelas VIII*, Intan Pariwara 2008. The summative test was aimed to eighth grade students at the first semester in the school. The data described in this study are taken from the students of MTsN-1 Model Palangka Raya class VIII-1. The numbers of the students are 36 students. The test was given on May 23<sup>rd</sup>, 2009 at MTsN-1 Model Palangka Raya. The test was done in allocated time of 90 minutes (11.30-13.00 pm). The description of the summative test can be seen on appendix.

##### **B. The Validity of English Summative Test Items of Multiple Choices Made by the English teachers of MTsN-1 Model Palangka Raya.**

Based on the previous chapter, it has been explained that the validity of a test is the extent to which the test measures what is intended to measure. If the test items are valid, so they can be used to measure the ability of the students.

There were some ways to analyze the validity of the test as follows:

### **1. Content Validity**

Content validity is also called curricular validity. The test is said has content validity if the content is proper with the scope and content of curriculum in school. Content test has suitable or represent the sample of learning result that should be achieved according to curriculum. To measure the validity of the test items, the writer used content validity by compared the curriculum materials used with the test items. So, look at the table (appendix 2) to see the appropriate the test and curriculum by seeing the curriculum and the handbook of English teacher.

From the table can be known that the content of curriculum and material in handbook is evenly distributed in the Multiple Choice test. It means every material that in curriculum has been representatives in multiple choice tests. A good test contents every material that has been taught. Based on the table we can say that the English multiple choices has validity content which means the English multiple choice has effectiveness based on the contents representativeness.

There are the analysis procedures of content validity:

- a) Preparing the curriculum and syllabus used by the teacher.
- b) Comparing the curriculum and syllabus material with the content of the English summative test items.
- c) Tabulating the result.

### **2. Construct Validity**

A test is said has construction validity if the item is arranged to be able to measure every aspect of thinking that is written in curriculum. In this case, the analysis of the



construction validity test has done by seeing the syllabus of English subject at MTsN-1 Model Palangkaraya which use competency based curriculum 2004. See on appendix 3 to know the appropriateness of the test with syllabus.

Based on the appendix, it can be seen that the item is not distributed evenly based on syllabus. There's one standard competence that not included in English Multiple Choice in MTsN-1 Model Palangkaraya that is listening and writing skill. Based on the table we can see there are 9 item for speaking (2, 7, 6, 13, 14, 22, 25, 12, 36), 10 for reading (8, 9, 20, 21, 32, 33, 34, 38, 39, 40), and no one for listening and for writing. So it can be concluded that the Multiple Choice items for English test in MTsN-1 Model Palangkaraya is enough representative.

Based on what had been stated above. It can be calculated:

$$= \frac{19}{40} \times 100\%$$

$$= 48\%$$

From the result of calculation above writer declare that the validity of English Multiple Choice items in MTsN-1 Model Palangka Raya at first semester is having fair validity.

### **C. The Reliability of English Summative Test Items of Multiple Choices Made by the English teachers of MTsN-1 Model Palangka Raya.**

The formula that is used to know the reliability of Multiple Choice test for English subject in MTsN-1 Model Palangkaraya by using Kuder Richardson method, which is known KR 20 and KR 21. In this writing, writer choosed KR 20 formula to know the reliability of the test that is presented by Sudijono as follow:

The writer uses KR-20 formula to analyze the reliability of the test:

$$r_{11} = \left( \frac{n}{n-1} \right) \left( \frac{St^2 - \sum p_i q_i}{St^2} \right)$$

Where:

$r_{11}$  : Coefficient reliability of test

$n$  : Number of test items.

1 : Constant multiplier.

$St^2$  : Total of Varian

$p_i$  : Mean of the correct answer.

$q_i$  : Mean of the wrong answer.

$\sum p_i q_i$  : Total multiplication between  $p_i$  and  $q_i$ .<sup>51</sup>

The qualifications of reliability are as follow:

$r$  between 0,00 – 0,20 : no correlation

0,20 – 0,40 : low

0,41 – 0,70 : fair

0,71 – 0,90 : high

0,91 – 1,00 : very high (perfect).<sup>52</sup>

---

<sup>51</sup>Anas Sudijono, *Pengantar Evaluasi Pendidikan*, Jakarta: PT. Raja Grafindo Persada, 2005. p. 252-253.

<sup>52</sup>Ngalm Purwanto, *Teknik Evaluasi*, p. 139.

Based on data in table 2, the following is the calculation to find out the test reliability. It is known:

$$X = 1253$$

$$\Sigma X^2 = 44111$$

$$\begin{aligned}\Sigma Xt^2 &= \Sigma X^2 - \left( \frac{(X)^2}{N} \right) \\ &= 44111 - \left( \frac{1253^2}{36} \right) \\ &= 44111 - 43611.4 \\ &= 499.6\end{aligned}$$

$$\begin{aligned}St^2 &= \left( \frac{\Sigma Xt^2}{N} \right) \\ &= \frac{499.6}{36} \\ &= 13.9\end{aligned}$$

Now, it is known:  $n = 40$ ,  $\Sigma p_i q_i = 3.689$ ,  $St^2 = 13.9$

$$\begin{aligned}r_{11} &= \left( \frac{40}{40-1} \right) \left( \frac{13.9-3.689}{13.9} \right) \\ &= \left( \frac{40}{39} \right) \left( \frac{10.211}{13.9} \right) \\ &= 1.03 \times 0.73 \\ &= \mathbf{0.75}\end{aligned}$$

The result of the calculation, so  $r = 0.75$  means that the English of Multiple Choice test item in MTsN-1 Model Palangka Raya on first semester was high reliable.

**D. The Discriminating Power of English Summative Test Items of Multiple Choice Made by the English teachers of MTsN-1 Model Palangka Raya.**

Item discrimination analysis determines how well each item discriminates between high and low level examinees, for each item in a test should help to separate the proficient subjects from those who lack the tested skills or learning. The usual method is to assume that an examinee's performance on the total test will provide a reasonably good indication of his level of achievement. Therefore, the test analysis will first separate the highest and lowest papers in terms of the total scores on the test. Then determine how the two groups did on each item, the discriminating items will be those answered correctly by more of the high group than the low group.

Then the result of each item is tabulated, next by using the tabulating it can calculate the discrimination power. The tabulating to determine item discrimination is not far different with item difficulty, so here the item difficulty and discrimination power calculated mutually.

Formulation to determine item discrimination as follows:

$$DP = \frac{U - L}{1/2T}$$

Which:

DP = discriminating power or item discrimination

N = numbers of sample

U = number of correct answer from upper group

L = number of correct answer from upper group

T = number of upper and lower group<sup>53</sup>

Criterion:

DP: 0, 00-0, 20 : low

DP: 0, 20-0, 40 : fair

DP: 0, 40-0, 70 : good

DP: 0, 70-1, 00 : very good

Based on the result of Discriminating Power (see on appendix 4), the average discrimination power is 0, 1. In general the English Multiple Choice tests for eighth grade students of MTsN-1 Model Palangkaraya can be turned out has low discrimination power. More clearly, there are 30 items that should be revised; they are item number 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 32, 33, 34, 36, 38 and 39. While the test items have good discrimination power are 10 items.

**Table 2**  
**The Percentage of Discriminating Power**

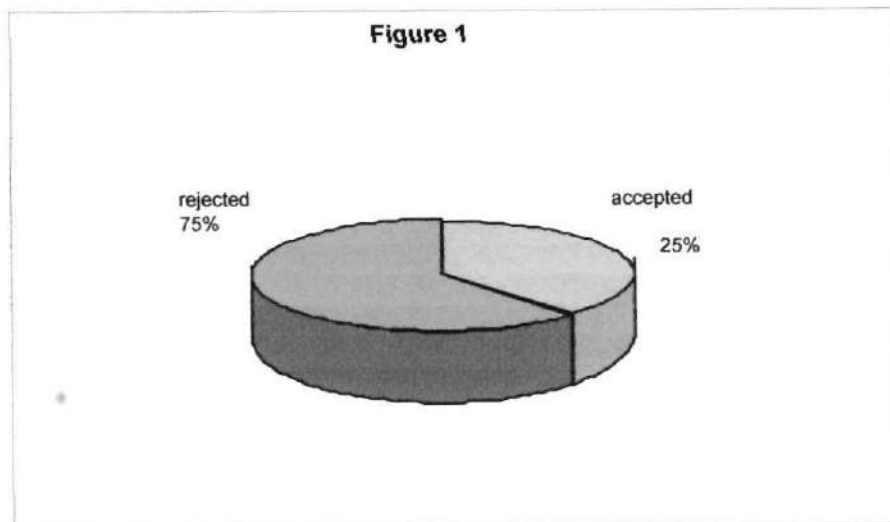
No		Item	%
1	Accept	10	25
2	Reject	30	75
			100

From the table above, it is known that there were 10 items (25%) accepted and there were 30 (75%) that should be refused.

---

<sup>53</sup>M. Ngalim. Purwanto, *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*. Bandung: PT. Remaja Rosdakarya, 2006, p. 120.

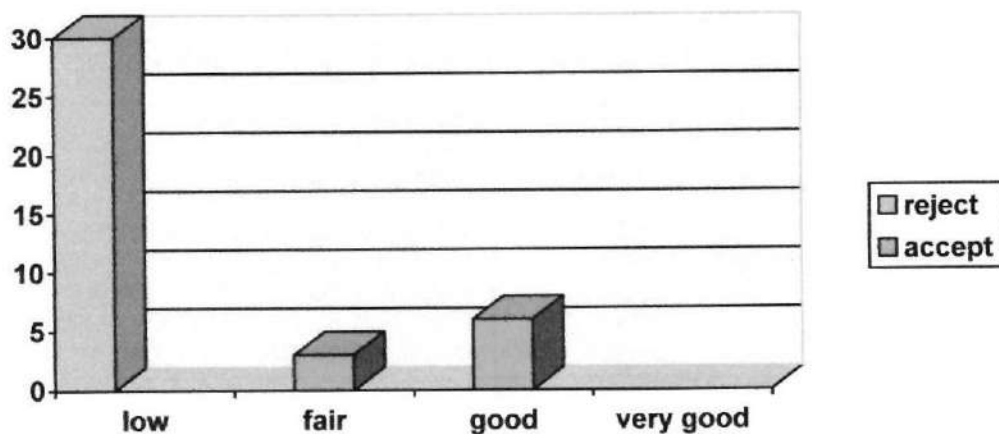
The description of percentage of discriminating power in figure below:



**The percentage of discriminating power**

**Chart 1**

**The frequency and percentage of discriminating power**



#### **E. The students' score of MTsN-1 Model in Summative test in academic 2008/2009**

The result of the students' score of MTsN-1 Model in Summative test could be seen in the table below:

**Table 3**

**The distribution of frequency and percentration of Students' scores in summative test items of Multiple Choice**

No	Scores	F	%
1.	31-40	30	83
2.	10-30	6	17
Total		36	100

Based on the table above, it can be explained that there were 30 (thirty) students (83%) who acquired score 31-40. Then, there were 6 (six) students (17%) who acquired score 10-30. it could be known that most the students obtained score 31-40, or it can be said that they obtained high score.

From the frequency distribution above, the writer constructed the chart as follows:

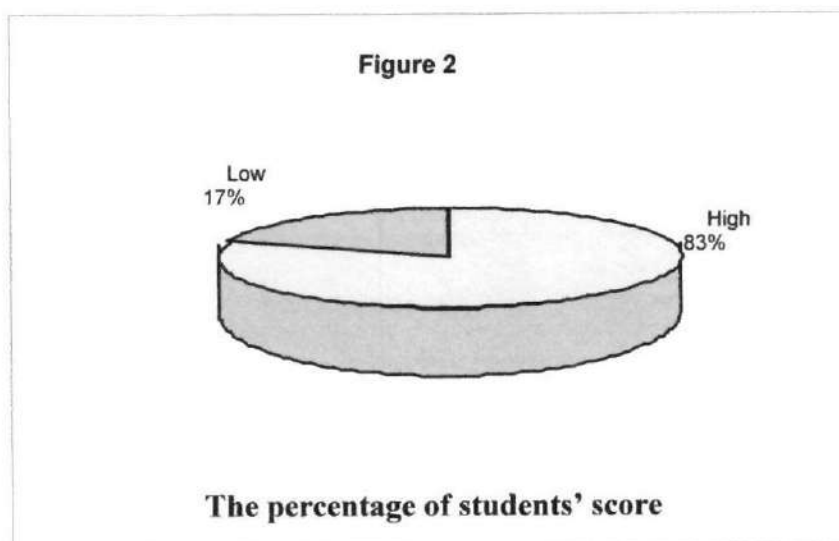
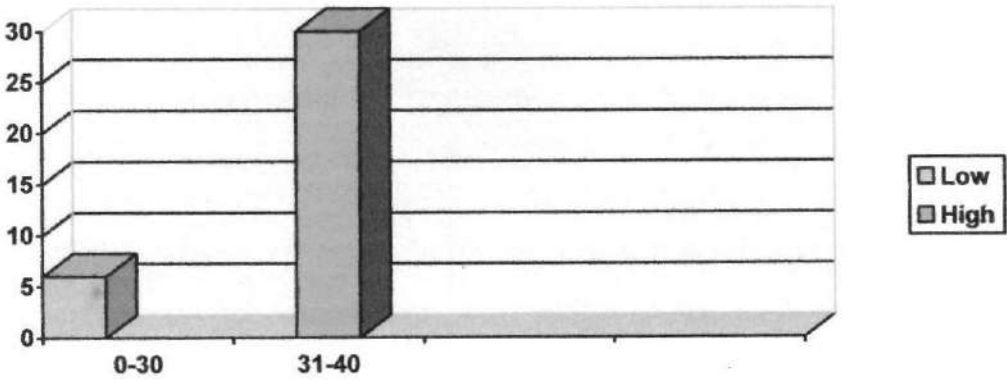


Chart 2

The frequency and percentation of student's score in summative test





# **CHAPTER IV**

## **CHAPTER IV**

### **CLOSING**

In this chapter, the writer concluded the facts based on the result of the study. Then, the writer tried to make some suggestion dealing with this study.

#### **A. Conclusion**

From the previous chapter, the writer presents the conclusions as follows:

The quality of English summative test items of Multiple Choice made by the English teachers at the eighth grade students in the first semester of MTsN-1 Model Palangka Raya. It was fair valid, low reliable and low discriminates. It shows on the result below:

##### **a. Validity**

###### **1) Content Validity**

Based on appendix 2, it shows English Multiple choice test in first semester at MTsN-1 Model Palangka Raya in academic year 2008/2009 is appropriate with the material teaching learning in handbook.

###### **2) Construct validity**

Based on appendix 3, it shows the English summative test items of Multiple Choice made by the English teacher at MTsN-1 Model Palangka Raya at first semester in academic year 2008/2009 doesn't quite representative.

###### **3) Validity test**

Based on finding result of validity test has shown that the score got by the English summative test items of Multiple Choice made by the English teachers at MTsN-1 Model Palangka Raya in first semester in academic years 2008/2009 are fair validity.

b. Reliability test

Based on finding result of validity test has shown that the reliability score got by English Multiple choice test in first semester summative test in academic year 2008/2009 at MTsN-1 Model Palangka Raya are 0.75 and it can be said high reliable.

c. Discriminating Power is known that there were 10 items (25%) accepted and there were 30 (75%) that should be refused.

**B. Suggestion**

Based on the finding fact of the study above, in this final section, there are several suggestions may be promoted as the following ways:

1. The English test items should be based on the curriculum and syllabus.
2. A teacher and all educational instruments should give pre-test the material and analyze the result of test to give feedback for the teacher is the test has already appropriate for students or not. Based on data gathered, the English summative test of multiple choice in MTsN-1 Model Palangka Raya for first semester in academic year 2008/2009 are low validity, reliability that is high reliable, and the discriminating is low.
3. The test that is made should not give final score based on the multiple choice test in summative test, because it is not quite effective to shows the real students mastery.
4. The English test items should be applied in try out and analyzed before they are given to the students.

## REFERENCES

- Arikunto, Suharsimi. 2003. *Dasar-Dasar Evaluasi Pendidikan*. Ed. Revisi. Cet.4. Jakarta: Bumi Aksara.
- , 2003, *Manajemen Penelitian*, Jakarta: Rineka Cipta.
- , 2006, *Prosedure Penelitian Suatu Pendekatan Praktik*, Jakarta: Rineka Cipta
- A.S.Hornby, 1995. *Oxford Advanced Learner's Dictionary of Current English*, Jonathan Crowther, Oxford: Oxford University Press.
- Cresswell, John. *Research Design Qualitative and Quantitative Approach*. California: SAGE Publication, inc.
- Departemen Pendidikan Nasional., 2003, *Kurikulum 2004 Sekolah Menengah Pertama (SMP)*, Jakarta.
- Gronlund, E. N. 1981. *Measurement And Evaluation In Teaching Fourth Edition*. Macmillan Publishing CO. INC. New York.
- Hadeli., 2006, *Metode penelitian Kependidikan*, Jakarta: PT. Ciputat Press.
- Hadma, Dina. 2008. The Validity and Reliability of Multiple Choice Test in Measuring Student's Achievement at MAN Model of Palangka Raya, *skripsi*, Palangka Raya: STAIN
- Halimi, Akhmad. 2008. Summative Evaluation Problems of English Lesson Faced by The Second Years Students of MTs-Muslimat NU Palangka Raya, *skripsi*, Palangka Raya: STAIN.
- Hasan, M. Iqbal. 2002. *Pokok-pokok Materi Metodologi Penelitian dan Aplikasinya*, Jakarta: Ghalia Indonesia.
- Heaton, J.B. 1974. *Writing English Language Test*. London: Longman Group Ltd.
- Nurdiyantoro Burhan, Gunawan, Marzuki. 2004. *Statistik Terapan*, cet. 3. Yogyakarta: Gadjah Mada University Press.
- Oktarini Melya Y.B., 2007. "The Quality of English Summative Test Made by The English Teacher of SMP Kristen Palangka Raya", *Skripsi*, Palangka Raya: Universitas Palangka Raya.
- Purwanto, Ngalim M. Drs. 2000. *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Cet.9. Bandung: Remaja Rosdakarya.

- Slameto, Drs. 1999. *Evaluasi Pendidikan*. Cet.2. Jakarta: Sinar Grafika Offset.
- STAIN Palangka Raya., 2007. *Pedoman Penulisan Skripsi*, Palangka Raya.
- Sudijono, Anas. 2005. *Pengantar Evaluasi Pendidikan*. Jakarta: Raja Grafindo.
- Sugiyono., 2004. *Statistika Untuk Penelitian*. Bandung: Alfabeta.
- , 2007. *Metode Penelitian Pendidikan: Pendekatan Kuantitatif, kualitatif, dan R&D*, Bandung: Alfabeta.
- Suharto., 1988. *Metodologi Penelitian Dalam Pendidikan bahasa; suatu pengantar*, Jakarta: Dediknas.
- Sukardi., 2003. *Metodologi penelitian pendidikan: Kompetensi dan Praktiknya*, Jakarta: PT. Bumi Aksara.
- Silverius, Suke. 1991. *Evaluasi Hasil Belajar dan Umpan Balik*. Jakarta: PT. Grafindo.
- Tinambunan, Wilmar Drs. 1988. *Evaluation of Student Achievement*. Departemen Pendidikan dan Kebudayaan.
- Thoha, M.C. 2003. *Teknik Evaluasi Pendidikan*. Ed. 1. Cet.5. Jakarta: Raja Grafindo Persada.
- Yulia, Rini. 2006. Kemampuan Guru Menentukan Nilai Akhir Mata Pelajaran PAI di SMA Palangka Raya, *skripsi*, Palangka Raya: STAIN.

www.park.edu.Com ( On line January 8<sup>th</sup> 2009)

<http://www.socialresearchmethods.net/kb/intreval.php> ( On line January 10<sup>th</sup> 2009)

<http://www.d.umn.edu/~schilton/Articles/GPD2.html> (On line December 2<sup>nd</sup> 2008)

[http://www.highlandschools-virtualib.org.uk/lit/whole\\_learner/summative.htm](http://www.highlandschools-virtualib.org.uk/lit/whole_learner/summative.htm) (On line November 18th 2008)

# APPENDIX

## Appendix 1

### THE STUDENTS' SCORE OF SECOND YEAR STUDENTS OF MTsN-1 MODEL PALANGKA RAYA

NO	CODES	TOTAL SCORES	X <sup>2</sup>
1	A1	37	1369
2	A2	35	1225
3	A3	37	1369
4	A4	28	784
5	A5	36	1296
6	A6	39	1521
7	A7	36	1296
8	A8	37	1369
9	A9	37	1369
10	A10	34	1156
11	A11	35	1225
12	A12	38	1444
13	A13	37	1369
14	A14	38	1444
15	A15	38	1444
16	A16	32	1024
17	A17	36	1296
18	A18	37	1369
19	A19	32	1024
20	A20	36	1296
21	A21	28	784
22	A22	38	1444
23	A23	24	576
24	A24	37	1369
25	A25	38	1444
26	A26	37	1369
27	A27	36	1296
28	A28	30	900
29	A29	35	1225
30	A30	33	1089
31	A31	32	1024
32	A32	39	1521
33	A33	38	1444
34	A34	28	784
35	A35	37	1369
36	A36	28	784
Total		1253	44111



Appendix 2

THE APPROPRIATE ITEM WITH CURRICULUM  
AND HANDBOOK MATERIAL

No	Material	Number of item	Item	Appropriate	
				Yes	No
1	Asking for, giving and refusing services	1,3,18	<ul style="list-style-type: none"> <li>➤ Rully : Do you need any help Zahra? Zahra : ..... What would Zahra say if she wants to refuse any help from Raffi? a. Thanks you b. Sorry, I can't c. Help me, please. d. No, really. I can manage, thanks.</li> <li>➤ Sinta says, "Would you take me a cup of coffee?". What does it mean? a. She is refusing service b. She is accepting service c. She is giving something d. She is asking for service</li> <li>➤ Dea : Could I carry your bag sir? Mr. Algif : .... I can do it by myself a. Yes please b. Sorry, I can't c. No, thanks d. With my pleasure</li> </ul>	√	
2	Asking for, giving and refusing things	4	<ul style="list-style-type: none"> <li>➤ The shop assistant say, "here is it". What does it mean? a. She is refusing service b. She is accepting service c. She is giving something d. She is asking for service</li> </ul>	√	
3	Descriptive	19	<ul style="list-style-type: none"> <li>➤ It is a kind of birth. It can imitate what someone says. It's a/an ... a. Eagle b. Parrot c. Pigeon d. Ostrich.</li> </ul>	√	
4	Announcements	11, 15-17, 23, 26	<ul style="list-style-type: none"> <li>➤ We usually find the notice above in a ... a. Class room b. Bus station c. Terminal d. Market</li> <li>➤ Can we read boks in the library on Sundays? a. Yes, we can b. Yes we will c. No, we can't d. No, we won't</li> <li>➤ How long is the open time of library on Saturday? a. 8 hours b. 6 hours c. 4 hours d. 2 hours</li> <li>➤ The sign means that we can ... here a. Stand b. Walk c. Park d. Work</li> <li>➤ Arrange the words below become a meaningful notice! You – not-for-thank-smoking 1    2    3    4    5  a. 4-3-1-2-5 b. 4-3-1-5-2 c. 4-3-2-5-1 d. 4-1-2-3-5</li> <li>➤ The public sign above is ...</li> </ul>	√	

			a. Instruction b. Warning c. Prohibition d. Information		
5	Lost and found ads	-	-		
6	Adjectives	27-30	TEXT ➤ a. famous b. small c. strategic d. original ➤ a. national b. local c. foreign d. original ➤ a. kinds b. ways c. types d. steps ➤ a. tour b. journey c. voyage d. flight	√	
7	Comparative and superlative adjectives	12	➤ In a book store. English book costs Rp.10.000, Math book Rp.12.500 and Indonesian book costs Rp.13.000. It means that Indonesian book is the ... book in the book store. a. most expensive b. more expensive c. cheapest d. cheaper	√	
8	The Simple Present Tense	38-40	TEXT : INDONESIA ARCHIPELAGO ➤ Where is Indonesia located? a. In Southeast Asia b. Near Mollucass c. Along equator d. In the Greenland ➤ Which of following belongs to the first largest island in the world? a. Borneo b. Irian Jaya c. New Guinea d. Greenland ➤ Which is the following statement is NOT TRUE according to the text? a. Most Indonesian population live in Irian Jaya b. Indonesia has more than 13,000 islands c. Borneo is the third biggest in the world d. The Mollucas is between Sulawesi and New Guinea	√	
9	Congratulation	2, 25	➤ Your best friend gets the highest score on the examination. What will you say? e. Happy Birthday f. I'm really sorry! g. How poor you are! h. I must congratulate you on your success. ➤ Today is 1 <sup>st</sup> January. What will you say to other people? a. Happy Idhul Adha b. Happy New Year c. Happy Birthday d. Merry Chrismast	√	
10	Complimenting	22	Rizka: Lk Meda! I get a new dress from my mother! Meda: <b>Wow, what a nice dress!</b> Rizka: Thank you ➤ The bold type words in the dialog is the expression of ... d. Complimenting e. Congratulation f. Asking for opinion g. Apologize	√	
11	Recount	8-9	➤ The kind of the text above is ... a. Recount	√	

			<ul style="list-style-type: none"> <li>b. Anecdote</li> <li>c. Narrative</li> <li>d. Descriptive</li> </ul> <p>➤ Here are characteristic of recount text, except ...</p> <ul style="list-style-type: none"> <li>a. Re-orientation</li> <li>b. Identification</li> <li>c. Orientation</li> <li>d. Even</li> </ul>		
12	Letters	20-21, 32-34	<p>Here are the functions of the letter, except ...</p> <ul style="list-style-type: none"> <li>a. to exchange information</li> <li>b. to maintain communication</li> <li>c. to give information</li> <li>d. to buy something</li> </ul> <p>➤ Palangka Raya, December 14, 2008. it's shows a part of letter about ...</p> <ul style="list-style-type: none"> <li>a. Date</li> <li>b. Body</li> <li>c. Greeting</li> <li>d. Closing</li> </ul> <p>LETTER</p> <p>In Andini's letter, she tells Anis about ...</p> <ul style="list-style-type: none"> <li>a. Her family and relative</li> <li>b. Her pet named Maniz</li> <li>c. Her favorite food</li> <li>d. Her activities in her spare time</li> </ul> <p>The text above is ...</p> <ul style="list-style-type: none"> <li>a. Descriptive</li> <li>b. Report</li> <li>c. Narrative</li> <li>d. Recount</li> </ul> <p>Based on the descriptions above, what is Maniz?</p> <ul style="list-style-type: none"> <li>a. a dog</li> <li>b. a cat</li> <li>c. a horse</li> <li>d. a hamster</li> </ul>	√	
13	The Simple Past Tense	10, 37	<p>➤ The students played basket ball three days ago. The correct interrogative form from the statement is ...</p> <ul style="list-style-type: none"> <li>a. Did the students play basket ball three days ago?</li> <li>b. Do students play basket ball three days ago?</li> <li>c. Did students played basket ball three days ago?</li> <li>d. Where the students played basket ball three days ago?</li> </ul> <p>➤ Mr. President ... Indonesian army in Lebanon two years ago.</p> <ul style="list-style-type: none"> <li>a. Visit</li> <li>b. Visits</li> <li>c. Visiting</li> <li>d. Visited</li> </ul>	√	
14	The Past Continuous tense				√
15	Approving and denying facts				√
16	Agreeing and disagreeing	6	<p>Amir : Good morning, Rusdi Rusdi : Morning Amir : How about going for swimming this afternoon? It's a nice day Tia : <u>That's a good idea. I agree with you</u></p> <p>The underline sentence expresses</p> <ul style="list-style-type: none"> <li>e. Disagreement</li> <li>f. Agreement</li> <li>g. Opinion</li> <li>h. Compliment</li> </ul>	√	
17	Asking for and giving opinions	7, 36	<p>Teacher: Eka, what is your opinion about English lesson? Eka : in my opinion English lesson is very difficult. I can't understand about the tense!</p> <p>➤ In the dialogue, the teacher is asking Eka's opinion about ...</p> <ul style="list-style-type: none"> <li>e. English lesson</li> <li>f. Eka's house</li> <li>g. The teacher's new dress</li> <li>h. Tenses</li> </ul>	√	

			a. Instruction b. Warning c. Prohibition d. Information		
5	Lost and found ads	-	-		
6	Adjectives	27-30	TEXT ➤ a. famous b. small c. strategic d. original ➤ a. national b. local c. foreign d. original ➤ a. kinds b. ways c. types d. steps ➤ a. tour b. journey c. voyage d. flight	√	
7	Comparative and superlative adjectives	12	➤ In a book store. English book costs Rp.10.000, Math book Rp.12.500 and Indonesian book costs Rp.13.000. It means that Indonesian book is the ... book in the book store. a. most expensive b. more expensive c. cheapest d. cheaper	√	
8	The Simple Present Tense	38-40	TEXT : INDONESIAN ARCHIPELAGO ➤ Where is Indonesia located? a. In Southeast Asia b. Near Mollucass c. Along equator d. In the Greenland ➤ Which of following belongs to the first largest island in the world? a. Borneo b. Irian Jaya c. New Guinea d. Greenland ➤ Which is the following statement is NOT TRUE according to the text? a. Most Indonesian population live in Irian Jaya b. Indonesia has more than 13,000 islands c. Borneo is the third biggest in the world d. The Mollucas is between Sulawesi and New Guinea.	√	
9	Congratulation	2, 25	➤ Your best friend gets the highest score on the examination. What will you say? e. Happy Birthday f. I'm really sorry! g. How poor you are! h. I must congratulate you on your success. ➤ Today is 1 <sup>st</sup> January. What will you say to other people? a. Happy Idul Adha b. Happy New Year c. Happy Birthday d. Merry Christmast	√	
10	Complimenting	22	Rizka: Lk Meda! I get a new dress from my mother! Meda: Wow, what a nice dress! Rizka: Thank you ➤ The bold type words in the dialog is the expression of ... d. Complimenting e. Congratulation f. Asking for opinion g. Apologize	√	
11	Recount	8-9	➤ The kind of the text above is ... a. Recount	√	

### Appendix 3

#### THE APPROPRIATE ITEM WITH THE SYLLABUS

No	Basic Competency	No Item	SMT	Item	Appropriate	
					Yes	No
1	<b>Listening</b> Response the meaning in transactional conversation to get things and interpersonal formally and informally using various simple spoken language accurately in daily lives and involves various speech act, such as : <ul style="list-style-type: none"> <li>• Merespon permintaan izin</li> <li>• Merespon permintaan persetujuan</li> <li>• Mengidentifikasi masalah</li> </ul>	-	1	-	-	-
	<b>Listening</b> Response the meaning accurately and acceptable in simple functional text (such as announcement, invitation) formally and informally in daily lives context, then also in descriptive and recount,			-	-	-
2	<b>Speaking</b> Expressing the meaning in transactional conversation to get things and interpersonal formally and informally using various simple spoken language accurately in daily lives and involves speech act such as : <ul style="list-style-type: none"> <li>• Meminta, memberi dan menolak memberi izin.</li> <li>• Meminta persetujuan, menyetujui, dan tidak menyetujui.</li> </ul>	6  13-14 35	1	Amir : Good morning, Rusdi Rusdi : Morning Amir : How about going for swimming this afternoon?  It's a nice day Tia : <u>That's a good idea. I agree with you</u> The underline sentence expresses a. Disagreement b. Agreement c. Opinion d. Compliment	√	
	<b>Speaking</b> Expressing the meaning accurately and acceptable in simple functional text (such as announcement, invitation) formally and informally in daily lives context, then also in descriptive and recount.			Ayu : Hello Vina. Vina : Hello Ayu Ayu : Would you like to come my birthday party this evening? Vina : I'd lve to. Thanks. Whats time is the party? Ayu : At 7.00 p.m From the dialog we know that ... a. Vina invites Ayu to attend her party; b. Ayu will come to Vina's party c. The party will be at 7 a.m d. Ayu asks Vina to come to her party ➤ The underline sentence is ... a. How to invite someone b. How to accept an invitation c. How to refuse an invitation	√	

	<p><b>Speaking</b> Expressing the meaning in transactional conversation to get things and interpersonal formally and informally using various simple spoken language accurately in daily lives and involves speech act such as :</p> <ul style="list-style-type: none"> <li>• Mengundang</li> <li>• Memuji</li> <li>• Memberi selamat</li> <li>• Mendeskripsikan seseorang</li> </ul>	22		<p>d. How to ask something ➤ Becham : Zidane, <b>would you accompany me to watch football competition?</b> Zidane : Sure, I would Becham : Ok, I will wait for you at 3.00 p.m The bold type of the dialogue is expression about ... a. Agreeing b. Disagreeing c. Asking for opinions d. Inviting</p> <p>Rizka: Lk Meda! I get a new dress from my mother! Meda: <b>Wow, what a nice dress!</b> Rizka: Thank you ➤ The bold type words in the dialog is the expression of ... a. Complimenting b. Congratulation c. Asking for opinion Apologize.</p>	✓	
		2		<p>➤ Your best friend gets the highest score on the examination. What will you say? a. Happy Birthday b. I'm really sorry! c. How poor you are! d. I must congratulate you on your success.</p>	✓	
		25		<p>➤ Today is 1<sup>st</sup> January. What will you say to other people? a. Happy Idul Adha b. Happy New Year c. Happy Birthday d. Merry Christmast</p>	✓	
	<p><b>Speaking</b> Expressing the meaning in transactional conversation to get things and interpersonal formally and informally using various simple spoken language accurately in daily lives and involves speech act such as :</p> <ul style="list-style-type: none"> <li>• Meminta, memberi dan menolak jasa</li> <li>• Meminta dan memberi pendapat</li> <li>• Mengidentifikasi masalah</li> </ul>	7, 36		<p>Teacher: Eka, what is your opinion about English lesson? Eka : in my opinion English lesson is very difficult. I can't understand about the tense! ➤ In the dialogue, the teacher is asking Eka's opinion about ... a. English lesson b. Eka's house c. The teacher's new dress d. Tenses ➤ Opick : What if we study in my house the afternoon? Marshanda:  I'm really sorry Opick : Never mind a. Certainly, I would b. No problem c. That's a good idea d. I'd love to, but I can't</p>	✓	
3	<p><b>Reading</b> Responding the meaning accurately and acceptable in simple written functional text (such as letter, undangan) formally and informally in daily lives context, then also in descriptive and recount.</p>	8-9	1	<p>➤ The kind of the text above is ... a. Recount b. Anecdote c. Narrative d. Descriptive ➤ Here are characteristic of recount text, except ... a. Re-orientation</p>	✓	

	<p><b>Reading</b> Responding meaning and rhetorical steps of essay written text in daily lives and to access knowledge in text in form: descriptive and recount</p>	<p>38-40</p>		<p>b. Identification c. Orientation d. Even</p> <p>TEXT : INDONESIAN ARCHIPELAGO</p> <p>➤ Where is Indonesia located?</p> <p>a. In Southeast Asia b. Near Mollucass c. Along equator d. In the Greenland</p> <p>➤ Which of following belongs to the first largest island in the world?</p> <p>a. Borneo b. Irian Jaya c. New Guinea d. Greenland</p> <p>➤ Which is the following statement is NOT TRUE according to the text?</p> <p>a. Most Indonesian population live in Irian Jaya b. Indonesia has more than 13,000 islands c. Borneo is the third biggest in the world d. The Mollucas is between Sulawesi and New Guinea.</p>	√	
		20-21		<p>Here are the functions of the letter, except ...</p> <p>a. to exchange information b. to maintain communication c. to give information d. to buy something</p> <p>➤ Palangka Raya, December 14, 2008. it's shows a part of letter about ...</p> <p>a. Date b. Body c. Greeting d. Closing</p>	√	
		32-34		<p>LETTER</p> <p>In Andini's letter, she tells Anis about ...</p> <p>a. Her family and relative b. Her pet named Maniz c. Her favorite food d. Her activities in her spare time</p> <p>The text above is ...</p> <p>a. Descriptive b. Report c. Narrative d. Recount</p>	√	
4	<p><b>Writing</b> Expressing the meaning in short functional text in formally and informally accurately, fluently, acceptable in context Expressing meaning and rhetorical steps accurately, fluently, and acceptable using various written language in daily lives context in descriptive and recount.</p>	-	1	-	-	-

## Appendix 4

### DISCRIMINATING POWER

Number of item	Num right answer (UP)	Num right answer (LOW)	Num all students	U+L	U-L	DP	Category
1	18	15	36	33	3	0.2	Low
2	18	18	36	36	0	0	Low
3	18	17	36	35	1	0.1	Low
4	18	14	36	32	4	0.2	Low
5	18	18	36	36	0	0	Low
6	18	17	36	35	1	0.1	Low
7	18	18	36	36	0	0	Low
8	18	18	36	36	0	0	Low
9	18	15	36	33	3	0.2	Low
10	17	16	36	33	1	0.1	Low
11	18	15	36	33	3	0.2	Low
12	18	16	36	34	2	0.1	Low
13	18	15	36	33	3	0.2	Low
14	18	9	36	27	9	0.5	Good
15	18	18	36	36	0	0	Low
16	18	13	36	31	5	0.3	Fair
17	18	17	36	35	1	0.1	Low
18	17	15	36	33	1	-0.1	Low
19	17	18	36	34	0	-0.1	Low
20	18	16	36	34	2	0.1	Low
21	18	18	36	36	0	0	Low
22	11	9	36	20	2	0.1	Low
23	17	14	36	31	3	0.2	Low
24	18	17	36	35	1	0.1	Low
25	18	18	36	36	0	0	Low
26	7	11	36	17	-5	-0.2	Low
27	18	11	36	29	7	0.4	Fair
28	17	10	36	27	7	0.4	Fair
29	16	11	36	26	6	0.3	Fair
30	16	6	36	22	10	0.6	Good
31	15	10	36	25	5	0.3	Fair
32	18	16	36	34	2	0.1	Low
33	18	17	36	35	1	0.1	Low
34	18	18	36	36	0	0	Low
35	18	10	36	28	8	0.4	Fair
36	18	15	36	33	3	0.2	Low
37	18	13	36	31	5	0.3	Fair
38	3	12	36	16	-8	-0.5	Low
39	18	14	36	32	4	0.2	Low
40	18	11	36	29	7	0.4	Fair
<b>Average Score</b>						<b>0.1</b>	<b>LOW</b>



## Appendix 5

### Key Answers

1) D	21) A
2) D	22) A
3) D	23) D
4) C	24) B
5) A	25) B
6) B	26) C
7) A	27) A
8) A	28) C
9) B	29) B
10) A	30) D
11) A	31) D
12) A	32) B
13) D	33) A
14) B	34) D
15) C	35) D
16) D	36) D
17) C	37) D
18) C	38) C
19) B	39) D
20) D	40) A

# Appendix 7

## TABLE OF ITEM ANALYSIS OF A TEST CONSIST OF 40 ITEMS

Code of students	Score																																								xt	Xt <sup>2</sup>	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40			
A1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	37	1369	
A2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	1	1	1	1	0	35	1225
A3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	37	1369
A4	0	1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	0	0	0	0	1	1	1	1	0	1	0	1	1	0	28	784	
A5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	36	1296	
A6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	39	1521	
A7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	36	1296
A8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	37	1369
A9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	37	1369
A10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1	1	1	1	0	1	1	0	1	0	34	1156	
A11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	1	1	1	1	0	35	1225
A12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	38	1444
A13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	37	1369
A14	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	38	1444
A15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	38	1444
A16	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	0	1	32	1024	
A17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	36	1296
A18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	37	1369	
A19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	0	1	0	0	0	1	1	1	0	1	0	1	1	32	1024		
A20	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	36	1296	
A21	0	1	1	0	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	0	0	1	1	1	1	1	0	1	0	1	1	0	28	784
A22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	38	1444
A23	1	1	0	0	1	1	1	1	0	0	1	0	0	0	1	1	0	1	1	0	1	0	1	0	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	0	0	24	576



Appendix 8

DISTRIBUTION OF UPPER GROUP

Code of students	Score																																								Score	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40		
A6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	39	
A32	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	39	
A12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	38	
A14	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	38	
A15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	38	
A22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	38	
A25	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	38	
A33	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	38	
A3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	37	
A8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	37	
A24	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	37
A1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	37
A9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	37	
A13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	37	
A18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	37	
A26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	37	
A35	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	37	
A5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	36	

Appendix 9

DISTRIBUTION OF LOWER GROUP

Code of students	Score																																								Score			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40				
A7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	36		
A17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	36	
A20	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	36		
A27	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	36		
A2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	1	1	1	1	0	35		
A10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1	1	1	1	0	1	1	0	1	0	35		
A11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	1	1	1	0	35		
A29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	0	1	1	1	1	0	1	0	1	1	35		
A30	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	0	1	33		
A16	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	0	1	32		
A19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	0	1	0	0	0	0	1	1	1	1	0	1	0	1	1	32	
A31	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	0	1	1	0	1	1	0	0	1	1	1	1	1	1	0	1	1	1	32			
A28	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	0	0	0	0	0	0	0	1	1	1	1	0	1	0	1	1	30
A4	0	1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	0	0	0	0	1	1	1	1	1	0	1	0	1	1	0	28		
A21	0	1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	0	0	0	0	1	1	1	1	1	0	1	0	1	1	0	28		
A34	1	1	1	1	1	1	1	1	0	0	1	0	1	0	1	0	1	1	1	0	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	0	1	0	1	0	1	28		
A36	0	1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	0	0	0	0	1	1	1	1	1	0	1	0	1	1	0	28		
A23	1	1	0	0	1	1	1	1	0	0	1	0	0	0	1	1	0	1	1	0	1	0	1	0	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	0	0	24		

THE NAME OF STUDENTS

NO	CODE	NAME
1	A1	AHMAD HUSEIN
2	A2	ALIA SYUKRIA AZIZAH
3	A3	ANISA FAJAR SURYANI
4	A4	ANNA ISTIQAMAH ZAHRA
5	A5	ANNISA
6	A6	ARINI AMIRAH
7	A7	AZKA MURTIA
8	A8	AYU PURNAMA SARI
9	A9	CAHYANI
10	A10	DEWITA FITRI
11	A11	FARHA KAMELIA
12	A12	FATHIR T.R
13	A13	FERDIRIGO SYAHRUL ASWAR
14	A14	HANAN S.
15	A15	HANIF MAULIDA ZUHRI
16	A16	HELDA PURNAMA SARI
17	A17	HINDON RAHMA WATI
18	A18	IJAI ABDUL KODIR GHANI
19	A19	IKRAMINA WULANSARI
20	A20	ISTI NORFAHRINA
21	A21	ISTIQMAH
22	A22	JOKO PRIYONO
23	A23	KURNIASARI
24	A24	MUHAMAD CHAIRIL RIZKYTA A.
25	A25	MUHAMMAD AAN MASLIANDY
26	A26	MUHAMMAD RIZKI NOPE YANDI
27	A27	MUHAMMAD AMIN SAYYAD
28	A28	NADYA AULIA RAHMI
29	A29	NINA ABADIAH
30	A30	NOORMI MAULIDINA
31	A31	NUR EKA INDAH SARI
32	A32	RAHMA ZAKIA
33	A33	RACHMAD MAULANA
34	A34	SUSI MUNAWARAH
35	A35	SYARIF HIDAYATULLAH
36	A36	WARDHA TILLAH

**PANITIA UJIAN SEMESTER GANJIL**  
**MTsN 1 MODEL PALANGKARAYA**  
**TAHUN PELAJARAN 2008/2009**

---

Mata Pelajaran	: Bahasa Inggris
Kelas/Semester	: VIII/I
Hari/Tanggal	: Kamis, 18 Desember 2008
Pukul	: 07.00-09.00 WIB

---

Choose a,b,c or d for correct answer!

1. Rully : Do you need any help Zahra?  
Zahra : .....  
What would Zahra say if she wants to refuse any help from Raffi?
- Thanks you
  - Sorry, I can't
  - Help me, please.
  - No, really. I can manage, thanks.
2. Your best friend gets the highest score on the examination.  
What will you say?
- Happy Birthday
  - I'm really sorry!
  - How poor you are!
  - I must congratulate you on your success

The dialog is for questions number 3 and 4

Santi : Can I help you?  
Sinta : Sure, Would you take me a cup of coffee  
Santi : Where is it?  
Sinta : in the kitchen, please.  
Santi : Here is it.  
Sinta : Thank you

3. Sinta says, "Would you take me a cup of coffee?". What does it mean?
- She is refusing service
  - She is accepting service
  - She is giving something
  - She is asking for service
4. The shop assistant say, "here is it". What does it mean?
- She is refusing service
  - She is accepting service
  - She is giving something
  - She is asking for service

5. Here are some expressions of invitations which one is correct to refuse invitations?

- a. I would love too, but I can't
- b. Sounds like good idea
- c. I'd like too.
- d. Okay

6. Amir : Good morning, Rusdi

Rusdi : Morning

Amir : How about going for swimming this afternoon? It's a nice day

Tia : That's a good idea. I agree with you

The underline sentence expresses

- a. Disagreement
- b. Agreement
- c. Opinion
- d. Compliment

7. Teacher : Eka, what is your opinion about English lesson?

Eka : in my opinion English lesson is very difficult. I can't understand about the tense!

In the dialogue, the teacher is asking Eka's opinion about ...

- a. English lesson
- b. Eka's house
- c. The teacher's new dress
- d. Tenses

This text for number 8 to 9

It was a sunny day. There was no body at home, because every body visited my grandma. I did not join them because my father asked me to stay at hme. I was lonely so I decided to walk along the street.

When I reached Arif's house there was somebody calling me. Arif called me from branch of a tree. He tried to caut a branch but he could nt. Then, I decided to help him cut a branch. At first, he doubled my ability, but I convinced him that im the expect. So, He gave the saw to me and I started to saw the branch.

I was a little bit careless and fell to the ground. Arif took me to the hospital to get medical treatment.

8. The kind of the text above is ...

- a. Recount
- b. Anecdote
- c. Narrative
- d. Descriptive

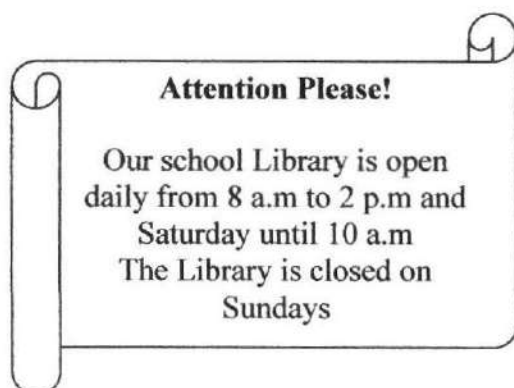
9. Here are characteristic of recount text, except ...


- a. Re-orientation
- b. Identification
- c. Orientation
- d. Even



10. The students played basket ball three days ago. The correct interrogative form from the statement is ...
- Did the students play basket ball three days ago?
  - Do students play basket ball three days ago?
  - Did students played basket ball three days ago?
  - Where the students played basket ball three days ago?
11. We usually find the ntice above in a ...
- Class room
  - Bus station
  - Terminal
  - Market
12. In a book store. English book costs Rp.10.000, Math bok Rp.12.500 and Indonesian book costs Rp.13.000  
It means that Indonesian book is the ... book in the book store.
- most expensive
  - more expensive
  - cheapest
  - cheaper
13. Ayu : Hello Vina.  
Vina : Hello Ayu  
Ayu : Would you like to come my birthday party this evening?  
Vina : I'd lve to. Thanks. Whats time is the party?  
Ayu : At 7.00 p.m  
From the dialog we know that ...
- Vina invites Ayu to attend her party
  - Ayu will come to Vina's party
  - The party will be at 7 a.m
  - Ayu asks Vina to come to her party
14. The underline sentence is ...
- How to invite someone
  - How to accept an invitation
  - How to refuse an invitation
  - How to ask something

For questions 15 to 16



15. Can we read books in the library on Sundays?
- Yes, we can
  - Yes we will
  - No, we can't
  - No, we won't
16. How long is the open time of library on Saturday?
- 8 hours
  - 6 hours
  - 4 hours
  - 2 hours
17. 
- The sign means that we can ... here
- Stand
  - Walk
  - Park
  - Work
18. Dea : Could I carry your bag sir?  
Mr. Algif : ... I can do it by myself
- Yes please
  - Sorry, I can't
  - No, thanks
  - With my pleasure
19. It is a kind of bird. It can imitate what someone says. It's a/an ...
- Eagle
  - Parrot
  - Pigeon
  - Ostrich
20. Here are the functions of the letter, except ...
- to exchange information
  - to maintain communication
  - to give information
  - to buy something
21. Palangka Raya, December 14, 2008. it's shows a part of letter about ...
- Date
  - Body
  - Greeting
  - Closing

22. Rizka : Lk Meda! I get a new dress from my mother!  
Meda : **Wow, what a nice dress!**  
Rizka : Thank you

The bold type words in the dialog is the expression of ...

- a. Complimenting
- b. Congratulation
- c. Asking for opinion
- d. Apologize

23. Arrange the words below become a meaningful notice!

You – not-for-thank-smoking

1      2      3      4      5

- a. 4-3-1-2-5
- b. 4-3-1-5-2
- c. 4-3-2-5-1
- d. 4-1-2-3-5

24. 1. It all happened very quickly, people just stood and watched.  
2. A car stopped outside the shop, and two men got out.  
3. He threw the brick at the window and smashed it.  
4. They had masked on  
5. One of them held a brick in his hand.  
6. One man took coats from from the window and the other put them in the back of the car.  
The best arrangement for the paragraph above is ...

- a. 2-3-4-1-6-5
- b. 2-4-5-3-6-1
- c. 2-4-6-3-1-5
- d. 3-4-5-6-2-1

25. Today is 1<sup>st</sup> January. What will you say to other people?

- a. Happy Idhul Adha
- b. Happy New Year
- c. Happy Birthday
- d. Merry Chrismast

26.

**DO NOT FEED  
THE ANIMAL**

The public sign above is ...

- a. Instruction
- b. Warning
- c. Prohibition
- d. Information

27. a. famous  
b. small  
c. strategic  
d. original
28. a. national  
b. local  
c. foreign  
d. original
29. a. kinds  
b. ways  
c. types  
d. steps
30. a. tour  
b. journey  
c. voyage  
d. flight
31. Rahma is ... She got 9 in her Test.  
a. good  
b. kind  
c. stupid  
d. generous

Letter for questions 32 to 34

Dear Anis,

In your previous letter you asked me about your pet. Here I'd like to tell you the kind.

You must guess what it is.

Well, my pet name is Maniz. It has four legs. It can run fast here and there inside the cage. It has soft white fur. It's so small and cute. I sometimes put it on my palm and stroke its fur. I feed it some vegetables

Can you guess what Maniz is?

I'm sure you can. It is a ...

Love  
Andini

32. In Andini's letter, she tells Anis about ...
- Her family and relative
  - Her pet named Maniz
  - Her favorite food
  - Her activities in her spare time
33. The text above is ...
- Descriptive
  - Report
  - Narrative
  - Recount
34. Based on the descriptions above, what is Maniz?
- a dog
  - a cat
  - a horse
  - a hamster
35. Becham : Zidane, **would you accompany me to watch football competition?**  
 Zidane : Sure, I would  
 Becham : Ok, I will wait for you at 3.00 p.m  
 The bold type of the dialogue is expression about ...
- Agreeing
  - Disagreeing
  - Asking for opinions
  - Inviting
36. Opick : What if we study in my house the afternoon?  
 Marshanda: \_\_\_\_\_  
 I'm really sorry  
 Opick : Never mind
- Certainly, I would
  - No problem
  - That's a good idea
  - I'd love to, but I can't
37. Mr. President ... Indonesian army in Lebanon two years ago.
- Visit
  - Visits
  - Visiting
  - Visited

The text is for questions number 38 to 40

#### Indonesian Archipelago

Indonesia is a country in Southeast Asia that consists of more than 13,000 islands. The islands lie along the equator and extend more than 5,000 kilometers. Many of the islands cover only a few square kilometers. But about a half of New Guinea and three quarters of Borneo also belong

to Indonesia. Both, islands are the second and the third largest islands in the world, after Greenland.

Borneo, Sulawesi, Java and Sumatra are included in the greater Sunda. The lesser Sunda islands extend from Bali eastward to the Timor. The Mollucass lie between Sulawesi and New Guinea. The western part of New Guinea is called Irian Jaya, and Indonesia territory. Compared to other regions, Irian Jaya is the most thinly populated.

38. Where is Indonesia located?
- a. In Southeast Asia
  - b. Near Mollucass
  - c. Along equator
  - d. In the Greenland
39. Which of following belongs to the first largest island in the world?
- a. Borneo
  - b. Irian Jaya
  - c. New Guinea
  - d. Greenland
40. Which is the following statement is NOT TRUE according to the text?
- a. Most Indonesian population live in Irian Jaya
  - b. Indonesia has more than 13,000 islands
  - c. Borneo is the third biggest in the world
  - d. The Mollucas is between Sulawesi and New Guinea.

**\*GOOD LUCK\***



Picture 1  
The Test Activities of Students of MTsN-1 Model Palangka Raya



Picture 2  
The Test Activities of Students of MTsN-1 Model Palangka Raya



**PANITIA SEMINAR PROPOSAL SKRIPSI MAHASISWA  
SEKOLAH TINGGI AGAMA ISLAM NEGERI  
STAIN PALANGKA RAYA**

Jl.G.Obos Komplek Islamic Center Tlp. (0536) 3239447/3226356 Fax. 3222105 Palangka Raya 73112

**SURAT KETERANGAN**

**No: 25 /PAN-SPSM/SG/III/2009**

Panitia Seminar Proposal Skripsi Mahasiswa Sekolah Tinggi Agama Islam Negeri (STAIN) Palangka Raya, menerangkan bahwa :

N a m a : NOOR AMALIA S.  
NIM : 030 112 0110  
Jurusan / Prodi : TARBIYAH/TBI  
Judul Proposal : THE QUALITY OF ENGLISH SUMMATIVE TEST ITEM OF  
MULTIPLE CHOICE MADE BY THE ENGLISH TEACHERS OF  
MTsN-1 MODEL PALANGKA RAYA

Telah melaksanakan Seminar Proposal Skripsi pada tanggal 7 Maret 2009 di Ruang Aula STAIN Palangka Raya dengan Penanggung Utama : **SABARUN,M.Pd** Moderator : **SYAIRIL FADLI,M.Hum** dan dinyatakan **lulus** dapat diterima sebagai syarat penyelesaian skripsi.

Palangka Raya, 13 Maret 2009

**PANITIA**

Ketua,

Sekretaris,

**ASMAWATI,M.Pd**

NIP. 150 311 460

**HARLES ANWAR,M.Si**

NIP.150 292 523





DEPARTEMEN AGAMA  
SEKOLAH TINGGI AGAMA ISLAM NEGERI  
(STAIN) PALANGKA RAYA

Alamat Jl. G. Obos Komplek Islamic Centre Telp. (0536) 39447, 26356, 21438 Fax 22105 Palangkaraya 73112

Palangka Raya, 2 April 2009

Nomor : Stt.15.8/TL.00/ /2009.  
Lampiran : 1 (Satu) Proposal.  
Perihal : Mohon Ijin Observasi /Penelitian.

Kepada

Yth. Kepala Kantor Departemen Agama  
Kota Palangka Raya  
di -  
PALANGKA RAYA

Sehubungan dengan salah satu tugas mahasiswa untuk mengakhiri studi pada Sekolah Tinggi Agama Islam Negeri (STAIN) Palangka Raya adalah membuat Skripsi, maka dengan ini kami mohon kiranya Bapak berkenan memberikan Izin Penelitian Lapangan kepada :

Nama : Noor Amalia Shanti  
NIM : 0301120110  
Jurusan/Prodi : Tarbiyah / TBI  
Jenjang : Strata 1 (S.1)  
Lokasi Penelitian : MTs N 1 Model Palangka Raya  
Judul Skripsi : "THE QUALITY OF ENGLISH SUMMATIVE TEST  
ITEM OF MULTIPLE CHOICE MADE BY THE  
ENGLISH TEACHERS OF MTsN-1 MODEL  
PALANGKA RAYA "  
Metode : Dokumentasi  
Waktu Pelaksanaan : 2 (dua) bulan, terhitung sejak tanggal 4 April s/d  
4 Juni 2009.

Sebagai bahan pertimbangan terlampir Proposal Penelitian, demikian atas perhatian dan pertimbangan Bapak disampaikan terima kasih.



An. Ketua

(Bantuan Ketua I)

BUBAKAR HM, M.Ag.  
NIP. 150 213 517.

Tembusan :

1. Yth. Ketua STAIN Palangka Raya (Sebagai Laporan)
2. Yth. Kepala MTsN 1 Model Palangka Raya
3. Arsip.



**DEPARTEMEN AGAMA**  
**KANTOR KOTA PALANGKARAYA**  
Jalan AIS Nasution PO Box 40 ☎ (0536 ) 3221968 Palangka Raya 73111

**REKOMENDASI**

Nomor : Kd.15.06 /4/ PP.00.9/653 / 2009

Berdasarkan surat Ketua Sekolah Tinggi Agama Islam Negeri (STAIN) Palangka Raya nomor : Sti.15.8/TL.00/ /2009 tanggal 2 April 2009 perihal mohon ijin observasi / penelitian, dengan ini Kepala Kantor Departemen Agama Kota Palangka Raya memberikan rekomendasi kepada :

N a m a : **Noor Amalia Shanti**  
NIM : 0301120110  
Jurusan / Prodi : Tarbiyah / TBI  
Jenjang : Strata - 1 (S-1)  
Lokasi Penelitian : MTsN 1 Model Palangka Raya  
Judul Skripsi : : *The Quality of English Summative Test Item of Multiple Choice of MTsN-1 Model Palangka Raya*  
  
Metode : dokumentasi  
Waktu Pelaksanaan : 2 ( dua ) bulan, terhitung sejak tanggal 6 April s.d. 6 Juni 2009

Untuk mengadakan Penelitian dengan ketentuan :

1. Segera melaporkan diri kepada Kepala Madrasah yang bersangkutan
2. Selama melaksanakan Penelitian tidak mengganggu Proses Belajar Mengajar ( PBM )
3. Setelah selesai melaksanakan Penelitian agar melaporkan hasilnya secara tertulis ke Kantor Departemen Agama Kota Palangka Raya Cq. Kasi Mapendais.

Demikian rekomendasi ini diberikan kepada yang bersangkutan untuk dipergunakan sebagaimana mestinya.

Dikeluarkan di: Palangka Raya  
Pada Tanggal : 6 April 2009



MASRANI ARSYAD  
NIP. 19550225 552

Tembusan :

1. Kakanwil Depag Prop. Kalteng Up. Kabid Mapendais di Palangka Raya;
2. Ketua STAIN Palangka Raya di Palangka Raya;
3. Kepala MTsN 1 Model Palangka Raya di Palangka Raya.



**DEPARTEMEN AGAMA**  
**MADRASAH TSANAWIYAH NEGERI 1 MODEL**  
**TERAKREDITASI "A"**  
Alamat : Jl. AIS Nasution No. 3 Telp. (0536) 22865  
**PALANGKA RAYA 73111**

SURAT KETERANGAN  
Nomor MTs.15.06.1/ 11.00/399/2009

Berdasarkan Surat Rekomendasi Kepala Kantor Departemen Agama Kota Palangka Raya No. Kd.15.06/4/PP.009/653/2009 tanggal 6 April 2009 perihal Rekomendasi, Kepala MTsN 1 Model Palangka Raya menerangkan bahwa :

Nama	: Noor Amalia Shanti
NIM	: 0301120110
Jenjang	: Strata 1 (S.1)
Lokasi Penelitian	: MTsN 1 Model Palangka Raya
Judul Skripsi	: "THE QUALITY OF ENGLISH SUMMATIVE TEST ITEM OF MULTIPLE CHOICE MADE BY THE ENGLISH TEACHERS MTsN 1 MODEL PALANGKA RAYA"

Telah melaksanakan penelitian di Madrasah Tsanawiyah Negeri-1 Model Palangka Raya pada tanggal 4 April s.d 4 Juni 2009.

Demikian surat keterangan ini diberikan kepada yang bersangkutan untuk dapat dipergunakan sebagaimana mestinya.

Palangka Raya, 27 Juni 2009  
Syamsuddin, S.Pd.I  
NIP.19610512 198503 1 003

Tembusan :

1. Kakanwil Depag Prov.Kalteng Up.Kabid Mapendais di Palangka Raya;
2. Kakandepag Kota Palangka Raya;
3. Ketua STAN Palangka Raya di Palangka Raya.

BERITA ACARA DAN CATATAN HASIL UJIAN SKRIPSI MAHASISWA  
JURUSAN TARBIYAH STAIN PALANGKA RAYA  
SEMESTER ..... TAHUN AKADEMIK .....

Pada Hari ini Kamis Tanggal 19 Bulan Nov Tahun 2009 telah




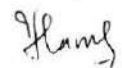
dilaksanakan ujian skripsi dengan judul :

The Quality of English Summative Test Item of Multiple Choice  
Made by the English Teachers of MTsN-1 Model Palangka Raya.

Nama : Noor Amalia Shanti  
Tempat/Tanggal Lahir : P. Raya, 3 April 1985  
N I M : 0301120110  
Jurusan/Prodi : Tarbiyah/...T.Pd

Tim Penguji :


Tanda Tangan

1. Drs. H. Abdul Gadir, M.Pd (  )
2. Ristati, M.Hum (  )
3. Dr. H. Khairil Anwar, M.Ag (  )
4. Ira. Halimah, M.Pd (  )

CATATAN HASIL UJIAN SKRIPSI

1. Check the parallelism, mechanics, choice of actions, redundant paragraphs
2. Revise the methodology into Quantitative : Conclusion & Suggestion  
Based on the problems of the study.
3. Put the hypothesis rejected clearly
4. Omit : instrument try out because it is about others' work, omit  
the population & sample ; test is the subject. ~ Content analysis.
5. ~~Ext~~ Use Exhaustive Conclusion : the students scores weren't good -  
because ....

Palangka Raya,  
Sekretaris

  
(.....)  
NIP. 19671226 199603 2003

## Appendix 6

### The English Teachers of MTsN-1 Model Palangka Raya

Name of Teachers
Akhmad Ali Mirza, S.Pd
John Syarif, S.Pd
Supardi, S.Pd
Yuli Dwi Astuti, S.Pd