AN ITEMS ANALYSIS OF ENGLISH FINAL TEST AT THE EIGHT GRADE OF MTs N KLATEN IN THE ACADEMIC YEAR OF 2016/2017

THESIS

Submitted as A Partial Requirements

For Undergraduate Degree in English Education Department

By:

ERNI TRI RAHAYU

SRN.133221077

ENGLISH EDUCATION DEPARTMENT

ISLAMIC EDUCATION AND TEACHER TRAINING FACULTY

THE STATE ISLAMIC INSTITUTE OF SURAKARTA

2018

ADVISOR SHEET

Subject   : Thesis of Erni Tri Rahayu

SRN   : 133221077

To :

The Dean of Islamic Education

and Teacher Faculty

IAIN Surakarta

In Surakarta

Assalamu'alaikum Wr.Wb.

After reading thotoughly and giving necessary advices, herewith, as the advisor, I statethat the thesis of

Name : Erni Tri Rahayu

SRN   : 133221077

Title   : "An Items Analysis of English Final Test at the Eighth Grade of MTs N Klaten in the Academic Year 2016/2017)"

has already fulfilled the requirements to be presented before The Board of Examiners (Munaqosyah) to gain the Undergraduate Degree in EnglishEducation Departement.

Thank you for the attention.

Wassalamu'alaikum Wr.Wb.

Surakarta, January 23th 2018

Advisor,

Novianni Anggraini, S.Pd, M.Pd

NIP. 19830130 201101 2 008

RATIFICATION

This thesis is to certify Undergraduate Degree thesis Entitled "AN ITEMS ANALYSIS OF THE ENGLIH FINAL TEST AT THE EIGHTH GRADE OF MTs N KLATEN IN THE ACADEMIC YEAR 2016/2017)" by Erni Tri Rahayu has been approved by the Board of Thesis Examiners as the requirement for the Undergraduate Degree Thesis in the State Islamic Institute of Surakarta.

Examiner I   : **Dr. Yusti Arini, M.Pd**   (.....................)

NIP. 19750829 200312 2 001

Examiner II   : **Dr. Hj, Woro Retnaningsih, M.Pd.**   (.....................)

NIP. 19681017 199303 2 002

Main Examiner   : **Novianni Anggraini, S.Pd, M.Pd**   (.....................)

NIP. 19830130 201101 2 008

Surakarta, 6 March 2018

Approved by

The Dean of Islamic Education and Teacher Training Faculty

Dr. H. Giyoto, M.Hum.

NIP. 196702242000031001

DEDICATION

This thesis is dedicated to:

1. My beloved parents who always support me until finished this thesis
2. My beloved sister, her husband and my nephews.
3. My big families.
4. My best friends Evi Wuria Santi and Anik Dwi K
5. All members of Crayon Class
6. IAIN Surakarta

MOTTO

Life is like riding bycycle.

To keep your balance, you must keep moving.

(Albert Einstein)

If you can't fly, then run .

If you can't run, then walk.

If you can't walk, then crawl.

But whatever you do,

You have to keep moving forward.

(Martin Luther King Jr.)

Barang siapa menolong Agamanya Allah maka Allah akan menolongnya dalam hal apapun, maka dahulukan Akhirat terlebih dahulu sebelum mengejar dunia mu

(Muhammad : 7)

PRONOUNCEMENT

Name   : Erni Tri Rahayu

SRN   : 133221077

Study Program   : English Education

Faculty   : Islamic Education and Teacer Training Faculty

I hereby sincerely state that the thesis titled **"AN ITEMS ANALYSIS OF THE ENGLIH FINAL TEST FOR EIGHTH GRADE OF MTs N KLATEN IN THE ACADEMIC YEAR 2016/2017"** is my real masterpiece. The things out of my masterpiece in this thesis are signed by citetio and reffered in the bibliography. If later proven that my thesis has discrepancies, I am willing to take the academic sanctions in the form of repealing my thesis and academic degree.

Surakarta,   February 2018

The researcher

Erni Tri Rahayu
SRN. 13.32.2.1.077

ACKNOWLEDGMENT

Alhamdulillah, all praises be to Allah, the single power, the Lord of the universe, master of the day of judgment, God all mighty, for all blessings and mercies so the researcher was able to finish this thesis entitled **"AN ITEMS ANALYSIS OF THE ENGLIH FINAL TEST AT THE EIGHTH GRADE OF MTs N KLATEN IN THE ACADEMIC YEAR 2016/2017".** Peace be upon Prophet Muhammad SAW, the great leader and good inspiration of world revolution.

The researcher is sure that this thesis would not be completed without the helps, supports, and suggestions from several sides. Thus, the researcher would like to express her deepest thanks to all of those who helped, supported, and suggested her during the process of writing this thesis. This goes to:

1. Dr. Mudhofir, S.Ag, M.Pd., the Rector of the State Islamic Institute of Surakarta
2. Dr. H. Giyoto, M.Hum., as the Dean of Islamic Education and Teacher Training Faculty of State ofIslamic Institute of Surakarta.
3. Dr. Imroatus Solikhah, M.Pd., as the Head of English Education Department of State Islamic Institute of Surakarta.
4. Novianni Anggraini, S.Pd, M.Pd. as the advisor for her time, advices, help, guidance, patience, suggestion, and corrections to revise the mistake during the entire process of writing this thesis.
5. All the lecturers and official employees of Islamic Education and Teacher Training Faculty.
6. The headmaster of MTs N Klaten, thanks for permission to conduct the research and collect the data.
7. The English teacher of MTs N Klaten who always kindly helps the researcher in conducting this research.
8. All the second grade students of MTs N Klaten for their participation to be the subject of this research.
9. Her beloved parents, Sutoyo and Sudalmi.Thanks for giving the best in her life.

The researcher realizes that this thesis is still far from being perfect. The researcher hopes that this thesis is useful for the researcher in particular and the readers in general.

Surakarta,   February 2018

The Researcher

Erni Tri Rahayu

TABLE OF CONTENT

CHAPTER IV: DATA DESCRIPTION, RESEARCH FINDING

AND DISCUSSION

CHAPTER V: CONCLUSION AND SUGGESTION

ABSTRACT

Erni Tri Rahayu.2018. *An Items Analysis of the Englih Final Test at the Eighth Grade Students of MTs N Klaten in the Academic Year 2016/2017*. Thesis. English Education Department. Islamic Education and Teacher Training Faculty.

Advisor   : Novianni Anggraini, S.Pd, M.Pd

Key Words   : Item Analysis, Test

Analyzing the test item is importan because it will give information for the teachers, students, on how to process the teaching and learning well. Beside that, the teacher will know the quality of the test item and the effectiveness of the teaching instruction. This research aimed to to determine the quality of the english final test for eighth grade students of MTs N Klaten in the academic year 2016/2017 in the term of difficulty level, discrimination index and distraction evaluation.

This research design is descriptive quantitative research. This study is descriptive because it describes the difficulty level, discrimitating power and distraction evaluation of the english final test for eighth grade students of MTs N Klaten while quantitative used for identifying the level of difficulty, the discriminating index and the effectiveness of distractors of English Final test items for the eighth grade students of MTs N Klaten in the2016/2017 academic year.The technique of collecting the data is using document analysis while the technique of analyzing the data is by using Arikunto's formula. The data taken from the test result of 100 students answer sheets of eightclasses by using random sampling.

The result of that there are 14 multiple choice items (35%) and 3 essay items (60%) that included in the category of good item, 15 multiple choice items (37.5%) and 2 essay items (40%) included as marginal items andc11 multiple choice items (27.5%) incuded as poor items.

LIST OF TABLE

LIST OF APPENDICES

CHAPTER I
INTRODUCTION

A. Background of the Study

Education is needed by everyone, it can be said that education is experienced by all people. Education have an important role in ensuring the survival of the nation and the state. Education as a means to improve the Human Resources (HR). Human Resources quality can bring the progress of Science and Technology in all aspects of life and bring people to a more advanced civilization and logical mindset (Rukmana, 2013). Human resource development can be enhanced by improving the quality of education in schools, which is determined by various factors. These factors are teachers, students, environment, infrastructure, learning time and learning process.

Human Resources quality will appear along with a good quality of education as well. Until recently, the education component that is used, needs to be improved, especially in the evaluation system of learning outcomes. Evaluation is a process to value something. To determine the value of the thing is being valued, it takes measurement, and the realization of the measurement is what commonly known in education called a test. Evaluation is very important due to the improvement of the quality of the learning program and wether the students reach the goals should be achieved (Sudijono: 2011).

One of form to evaluate the students' ability is test. Thus test could be a teacher-made test or standardized test. In the teacher-made test, the teacher who make the test should know and mastering the principles and the steps that must be done in making the test. By that knowledge, the teacher will get a clear figure about the general systematic framework of evaluation.

Testing is an important part of every teaching and learning. It contributes directly to the teaching learning process used in the classroom. Linn and Gronlund (200: 32) define tests as an instrument or systematic procedure for measuring a sample of behaviour by posing a set of questions

in the uniform manner. Based on that, test also included as a form of assesment which can be measure how well the students' achieve their language learning.

Teacher must do an evaluation after the test. It is also emphasized by the following quotation from NBPTS (National Board of Professional Teaching Standards):

> "Teacher are responsible for managing and monitoring students' learning. The teachers intend to recognize, know how to create, enrich, maintain, and alter instructional settings to capture and susrain students interest. They use many methods to measure students growth and understanding (Linn and Gronlund, 2000: 35)."

From those statement, the teachers are responsible to do the evaluation to establish how the teaching process can improve the students' achievement in the instructional objective. As the result we know the quality of the teacher's instruction and the students' ability. The teacher can use this information to make an improvement for further learning. Therefore, the teacher can increase the quality of learning process and learning outcome as a part of the increasing the quality of education system. Harmer (1998: 22) argues that a good teacher keeps a record of their students' achievements so that they are always aware of how they are getting on. By doing this, they can begin to see when teaching and learning has or has not been successful.

There are numerous types of test. There are placement test, achievement test, proficiency test and aptitude test. The test which is usually used by teacher to know how far students have mastered the lessons is the achievement test. The achievement test is intended to establish how successful individual students groups of students or the courses themselves have been achieving objectives of language courses. Then here are two kinds of achievement test: progress achievement test and final achievement test. Progress achievements are those intended to measure the progress that students are making and final achievement tests or summative tests are

intended to measure the students' achievement at the end of a course of study (Hughes, 1995: 10).

The test is a procedure that can be used to determine or measure something in accordance to the way and the rules that have been set, while the non-test is a procedure used to measure the affective domain such as attitudes, interests, talents, and motivation, for example using questionnaires, interviews, observations, and others (Sudijono, 2012: 67). Although there are two kinds of measuring instruments evaluation activities, but the test is often used for the evaluation tool. The test in question is a test that is made by the teacher (teacher-made test). The results of the test should reflect the real situation, because the results of the test will be used to make decisions. The size of the mistake (error) could affect the measurement of learning outcomes assessment, the test would be said as a good questions if they meet the requirements like eligibility, level of difficulty, distinguishing index, the pattern of answer distribution and relationship or correlation of each item with an overall score. Besides, the test should also have the characteristic of validity, reliability and objectivity. To find those things, it is needed an evaluation activity through the test item analysis to obtain information about the test which has been meets the requirements of a good question. Tests are said to be good as a measurement tool if they meet the requirements of the test, which has: validity, reliability, practicability, objectivity and economical (Suharsimi, 2013: 72).

According to Suharsimi (2013: 73), the data can be said to be valid if it is in accordance with the actual situation. If the data is valid, the instrument used is valid, because it is able to provide a description of the data correctly according to the actual situation. Test is said to be reliable if the test has been given repeatedly and provide consistent results (Brown, 2003). If there is no subjectivity factor in the scoring system, the test can be said as objective. Tests were carried out should be practical and easy for its administration. The test is easy to implement, easy to correct, and guided by clear instructions that can be given or initiated by others. The test is said to be economical if the implementation does not require high cost, a lot of labour, and a long time

(Suharsimi, 2013: 77). Tests that already meets some of the requirements that have been set, will be used as a measurement tool of student's achievement and learning success.

Through the test item analysis, it will be obtained information about the good and not good items. Good items will be retained and stored in the question bank, while the not good items should not be used again in the next assessment. The activity of revised the test items were performed in order to make the test qualified enough to be use as a measurement tools of student's learning outcomes. The test item analysis is done by calculating the aspect of validity, reliability, level of difficulty, distinguishing power, and the effectiveness of distractors (Zainal Arifin, 2011: 221).

Final test categorized as summative test Summative test aims to measure or summarize what a students grasped and it typically occurs at the end of the course or unit of instruction (Brown, 2003: 6). The importance of athe final test are; 1) monitoring the students' mastery of the learning maaterials, 2) motivating the students to study hard, 3) determining the semester grade. The research was held in MTs N Klaten which has moderate criteria. School Grade of *Kriteria Ketuntasan Minimal* (KKM) for English subject is 75 but based on the interview the score do not fulfill yet by the students.

Based on the interview toward some English teacher in MTs N Klaten, the test for the second semester of the eight grade students in the 2016/2017 academic year was made by MGMP (*Musyawarah Guru Mata Pelajaran)* of English subject. After the test was examined to the students, teacher do not analyze the English final test for the second grade students in the 2016/ 2017 academic year from quantitative aspect, so yet obtained the drawn about the quality of the test from quantitative and learning achievement which have done do not known yet. The limited ability of the teachers in MTs N Klaten in conducting test item analysis mostly caused by the lack of teachers' knowledge in conducting the test item analysis and the limitations of time and energy. The activity of test item analysis takes a lot of time and considerable effort, thus that teachers often feel burdened to do the

aatest item analysis. Eventhough, the activity of test item analysis is important to be concucted by the teacher to know the quality of the item test.

Thus far English teacher know about the difficulty level from the number of students who are able to answer that question. In preparing the test, teachers tend to follow the previous test. This condition will cause the tools of the test cannot show its quality as an evaluation tool of students in terms of level of difficulty, discriminating index, and the the distraction evaluation. This is why the english final test for the second grade students of MTs N Klaten in the 2016/ 2017 academic year needs to be analyzed.

Based on the explanation above, researcher interested to carry out the research entitled: "An Items Analysis of English Final Test at the Eighth Grade of MTs N Klaten in the Academic Year 2016/2017".

B. Identification of the Study

From the background study above, it can be identified several problems such as:

1. The lack of teachers' understanding about the technique of test item analysis to determine the quality of the questions.
2. Teachers feel burdened to analyze items because the process of analysis were complicated and take a long time.
3. The teacher know wether the test were good or not based on the number of answer and level of difficulty of the questions.

C. Limitation of the Problem

The scope of this population are too large. It is nearly impossible to study all of this problem. For that reason, the problem are limited on the Items Analysis of English Final Test for the Second Grade Students of MTs N Klaten especially in the term of difficulty level, discrimination index and distraction evaluation

D. Problem Statement

Based on the background described above, the main problem in this study are formulated as follows:

1. How is the difficulty level of the English Final Test for eighth grade students of MTs N Klaten in the academic year 2016/2017?

2. How is the discrimination index of the English Final Test for eighth grade students of MTs N Klaten in the academic year 2016/2017?

3. How is the effectiveness of the distractor of the English Final Test for eighth grade students of MTs N Klaten in the academic year 2016/2017?

E. The Objectives of the Study

The purpose of this study was to determine the quality of the english final test for eighth grade students of MTs N Klaten in the academic year 2016/2017 which cover:

1. Difficulty level

2. Discrimination index

3. Effectiveness of the distractor

F. The Benefits of the Study

The result of this study is hoped to give benefits theoritically and practically.

1. Theoritically Benefit

The result of this research will give a contribution to the development of educational world particularly to the process of item analysis of english final test for second grade students of MTs N Klaten, especially in the term of difficulty level, discriminating index and distraction evaluation.

2. Practical Benefit

This research is expected to be able to contribute some benefits to the researcher, English teacher and other research.

a. For the researcher, know the items analysis of the english final test for second grade students of MTs N Klaten, especially in the term of difficulty level, discriminating index and distraction evaluation.

b. For the English teachers, always developing good quality of the test.

c. For further research, it is expected that findings will be used as a starting points to conduct another research seeing as plenty of characteristics to make a good quality of the test.

G. Definition of Key Terms

1. Items Analysis

   According to Anthony (1983: 284), item analysis refers to the process of collecting, summarizing, and using information about individual test items especially information about pupil's response to items.

2. Test

   Test is a particular type of assessment that typically consists of a set questions administered during a fixed period of time under reasonably comparable conditions for all students (Linn &Gronlund, 1995: 5).

CHAPTER II

REVIEW ON RELATED LITERATURE

D. Theoritical Description

1. Theory of Teaching and Learning English

a. Definition of Teaching

Teaching is showing or helping someone to learn how to do something, giving instruction, guiding in the study of something, providing with knowledge, causing to know or understand. Teaching is guiding and facilitating learning, enabling the learner to learn, setting the conditions for learning. Teaching is complex process that can be conceptualized in many different ways, using alternative, metaphor, and analogies. Sujana states, teaching is to lead in student learning process, teaching is to regulate and organize the society in students environment, so that can forward and growth the student in their learning process.

Teaching means giving information which is the science and the skill to students which planed by systematically to match with the materials which curtained teaching is the creation a system of society which enable in teaching learning process. Teaching is the process of guiding pupil learning. In preparing any teaching procedure therefore the primary concern is : what are the steps of the learning procedure to be guided. These steps of the teaching formula and the teacher's must do what seems expedient to make learning satisfactory.

From the explanation above, we can find the word guiding, helping and also facilitating, in general it can be said that teaching is a kind of process how to make the students knowledge and attitudes improved.

b. Definition of Learning

According to Mouris, a learning is a systematic integrated outlook in regard to the nature of the process whereby people relate to their environments in such a way as to enhance their ability to use both themselves and their environments more effectively. Learning is a process which signed by variation in self individual. The variation is as the result of learning process and can shown in many forms, as like change in knowledge, understanding attitude and behavior, skill, capability, and their potential also. Change the other aspect in individual which learning. According to Cronbach, learning is shown by a change in behavior as a result of experience. Harold spears give definition, learning is to observe, to read, to imitate, to try something themselves, to listen, to follow direction.

Geoch said learning is a change in performance a result of practice. From definition above, we can conclude if a learning is change attitude or behavior, with, the junction of activity. For example by reading, listening, looking, following and other. Then, teaching and learning is has a fixed connection. We separate these process if for particular focus, study or descriptive action.

c. English Teaching and learning

Teaching and learning process is a kind of process in education that involves interactions from some components. They are the instructional objective the material, methods, media and evaluation as instruments to measure whether or not the process have a significant meaning, not only the interaction but also includes the educational aspect. The teacher's also have duty to transmit the values and skill so that they may improve their behavior and motivation in studying the materials.

The teacher occupies an important position in every learning teaching condition. Not only do they need to understand all the factors

with which they must deal, but he needs to know the ways in which these different factors can and should be handed to bring about the desired result. Teacher's activity as the term implies, includes everything the teacher does. It may enumerate these as :

a) electing unit and daily objectives
b) Selecting appropriate subjects matter and materials of instruction suited to the attainment of the objectives
c) Selecting pupil activities essential and desire able in the attainment of the objectives
d) Planning the organization of the classroom activities (motivation)
e) Securing the desired pupil activities
f) Directing the pupil activities
g) Evaluating the pupil activities
h) Evaluating the outcomes of pupil activity to determine the approximation of the desired outcomes (objectives).

2. Definition of Test

The importance of test has become a crucial milestone in the journey to individual's success. Test is a particular type of assessment that typically consists of a set questions administered during a fixed period of time under reasonably comparable conditions for all students (Linn & Gronlund, 1995: 5). Furthermore, according to Brown (2003: 3), test is a method of measuring a persons' ability, knowledge, or performance in a given domain. In order to know how well the result of learning process, teacher should evaluate it. By evaluating, teachers can collect information or have concept whether the teaching and learning activity has successes or not.

Gronlund said that "Tests are used as a means to motivate students to learn or review specific material" (Gronlund, 1982: 6). It means that test is one motivation of students to learn or review material in their school. Talking about test, it cannot be separated from assesment.

We might be attemped to think that "testing" and "assessing" are synonymous, but they are not. According to Linn and Gronlund (2000: 31) assessment is a broad term which covers the full range of procedures used to obtain information about the students' learning through observation, ratings of perfomances or projects and paper-and-pencil tests. It also covers the value judgments formation focusing the learning process. while, a test is a specific type of assessment which typicaly consists of a set of questions conducted during a fixed period of time under reasonably comparable condition for all students.

Based on the definitions above, the researcher concludes that the test is the particular types of assessment to reinforce learning and to motivate the students by giving a task or a set of tasks. Through the test, teachers don't only measure and motivate the students' ability but also improve the lesson in teaching learning process. In order to make a proper decision, the teacher needs an accurate data. So a good instrument is needed.

3. Types of Test

There are some types of tests in measuring students' learning achievement. Different experts give different classification of those tests. Those types of tests will be described bellow:

Linn and Grondlund (2000: 40) state that tests can be classified in the terms of their functional role in classroom interaction. It can be described as follows:

a. Placement test

A placement test is designed to determine pupil performance in the beginning of instruction.

b. Formative test

Formative test is intended to monitor learning progress during the instruction and to provide continuous feedback to both pupil and teacher concerning learning successes and failures. It is used at the end of a unit in the course book or after a lesson designed. The result of this test will give the students immediate feedback.

c. Diagnostic test

Diagnostic test is intended to diagnose learning difficulties during instruction. The main aim of diagnostic test is to determine the causes of learning difficulties and then to formulate a plan for remedial action.

d. Summative test

Summative test is a test which designed to assess students' achievement at the end of instruction. It is used to determine the extent to which the instructional goals have been achieved and to assign the course grade.

Summative test has clearly related to summative assessment. Summative assessment aims to measure or summarize what a student grasped and typically occurs at the end of a course or unit of instruction. A summation of what a student has learned implies looking back and taking stock of how well that student has accomplished objectives. But it does not necessarily point the way to future progress. Final exams in a course and general exams are examples of summative assessment (Brown, 2003: 06).

According to Brown, tests are divided into three categories there are achievement test, aptitude test, and proficiency

1) Achievement test

An achievement test is a systematic procedure for determining the amount a student has learned. Although the emphasis is on measuring learning outcomes, it should not be implied that testing is to be done only at the end of instruction (Gronlund, 1982: 1). While, Sudijonos' opinion (1996: 73), achievement test is test which is used to reveal the level of attainment or learning achievement. It is usually a formal examination given at the end of the school year or at the end of the course, the achievement test may be written and administered by

ministries of education, officially examining boards, or by members of teaching institution.

In order to have a good achievement test form, a test maker should consider that achievement test much be constructed well by paying attention to some following basic principles (Gronlund 1988: 303). They are:

a) Achievement tests should measure clearly defined learning outcomes that are in harmony with the instructional objectives.

b) Achievement tests should measure an adequate sample of the learning outcomes and subjects matter content included in instructions.

c) Achievement tests should include of the tests items, which are most appropriate for measuring the desired learning outcomes.

d) Achievement tests should be designed to fit the particular uses to be made of the results.

e) Achievement tests should be made as reliable as possible and should then be interpreted with caution.

f) Achievement tests should be used to improve student learning.

2) Aptitude test

According to Sudijono (1996: 73), the aptitude test is test which is executed that aim to reveal a basic competence or special aptitude that students have. Beside it, Brown states, a language aptitude test is designed to measures a persons' capacity or general ability to learn a foreign language and to be successful in that undertaking.

3) Proficiency test

The third type of test is proficiency test. This test is used to know the proficiency of test-takers. It is hoped after giving this test the test-taker will know their ability in their ability in language especially in English language. According to Hughes (2003: 11), proficiency tests are designed to measure people's ability in a language. The content of proficiency test is based on a

specification have to be able to do in the language in order to be considered proficient.

Based on the explanations about the kind of tests above, the researcherconcludes that generally test is a systematic and objective procedure to find out the knowledge and ability of what have been learned from someone.

E. Item Analysis

3. Definition of item analysis

According to Anthony, item analysis refers to the process of collecting, summarizing and using information about individual test items, especially information about pupils' responses to items. Analyzing the test item is importan because it will give information for the teachers, students, on how to process the teaching and learning well. Beside that, the teacher will know the quality of the test item and the effectiveness of the teaching instruction.

Item analysis usually concentrate three vital feature: level of difficulty, discriminating index and the effectiveness of each alternatives. Thus, item analysis information can tell us if an item was too difficult or too easy, how well the discriminate between high and low score of the test and whether all alternatives functioned as intended.

4. Kind of Item Analysis

a. Difficulty Level

There are many other names of difficulty level, those are: item facility, item difficulty, p-value etc. In the level of difficulty, there are three kind of difficulties, it from easy, moderate and difficult item. Item difficulty, which helps us decide if the test items are at the right level for the target group. Item facility expresses the proportion of the people taking the test who got a given item right.

According to Arikunto (1995: 211), Item difficulty is sometimes used to express similar information, in this case the proportion that got an item wrong. Where the test purpose is to make distinctions between candidates, to spread them out in terms of their performance on the

test, the items should be neither too easy nor too difficult. Good test is items which not too easy or not too difficult. If the items are too easy, then people with differing levels of ability or knowledge will all get them right, and the differences in ability or knowledge will not revealed by the item. Similarly if the items are too hard, then able and less able candidates alike will get them wrong and the item will not help us in distinguishing between them.

b. Item Discrimination

According to Arikunto (1995: 215), analysis of item discrimination addresses a different target: consistency of performance by candidates across items. The usual method for calculating item discrimination involves comparing performance on each item by different groups of test takers: those who have done relatively poorly. For example, as items get harder, we would expect those who do best on the vest overall to be ones who in the main get they right. Poor item discrimination indices are signal that an item deserves revision. If there are a lot of items with problems of discrimination, the information coming out of the test is confusing, as it means that some items are suggesting certain candidates that relatively better, while order individuals are better, no clear picture of the candidates' abilities emerges from the test.

c. The effectiveness of distractor

Arifin (2012: 282) states that the distractor effective when:

1) The total of students in lower and upper group who choice the distractor not less than:

$$25\% \times \frac{1}{2\sum d} \ x\,(ka + kb)$$

In which:

d = number of distraction

ka = upper group students

kb = lower group students

2) The distracters are chosen by the students in lower group instead of the students in upper group

3) Omit is not more than 10 % of the students.

Actually, the test item which can be called good quality that is distractor will be chosen by the students who answer incorrect equally. On the contrary, a poor test item is the distractor will be chosen enequally.

F. Previous Related Study

There are some studies about item analysis which had conducted by several researcher. There are three previous researches related to this research, as follow:

1. Arin Eka Primadani's Research

The first research is taken from journal article, written by Arin Eka Primadani (2013) under title "An Analysis of a Midterm English Test of the Seventh Grade Accelerated Class at SMPN 3 Malang". This study was conducted to evaluate the quality of the teacher-made test of the midterm English test of the 7 grade accelerated class at SMPN 3 Malang. This research was the qualitative and quantitative method, the quality of the test was analyzed based on the content validity, test reliability, item discrimination, item validity, and the effectiveness of the distracters. The sources of the data were the school documents and the students' score. The school documents were the test and the blueprint made by the teacher to develop the test.

Based on the analysis of the data, the test is considered to have low quality of the items in terms of the difficulty level, item discrimination, item validity, and the distracters. In contrast, there is no logical relationship between the results of content validity analysis with the item analysis of each item. The result of content validity analysis shows that the test items are 100% valid in covering the materials presented in the curriculum, but the results of the item analysis show that most of the items are not good. This case appears since there is no expert involved in this study to verify the content validity analysis.

2. Dian Nurlianto's Research

Dian Nurlianto did an analysis about validity, reliability, level of difficulty, distinguishing index, and the pattern of answer distributions of the Odd Semester Final Examination Questions in Accounting Economics Subject in Grade XII IPS SMA Negeri Banyumas Academic Year of 2014/2015. This study was included in descriptive quantitative research because the researcher used some numerical data which were analyzed statistically. In this study the researcher took 112 students as a sample.

The findings of this study were (1) items of multiple choice questions that can be said as valid totaling 27 items (77%), while the invalid items amounted to 8 items (23%). Whereas in the form of essay questions, all 5 items are categorized as valid (100%); (2) items of multiple choice questions can be said as reliable because the r lower than 0.70 or equal to 0.66. While the essay questions have lower than 0,70 or equal to 0.52 so that the items can be said as unreliable; (3) items of multiple choice questions that categorized to easy items are 19 items (52.3%), medium category 11 items (31.4%), and hard category are 5 items (14.3%).

While in essay questions, there are 2 items (20%) in easy category and 3 items (60%) in medium category; (4) items of multiple choice questions that categorized to poor items are 10 items (28.6%), satisfactory category 19 items (54.3%), and good category are 6 items (17.1%). While in essay questions, there is 1 item (20%) in poor category, 1 item (20%) in satisfactory category and 3 items (60%) in good category; (5) items (14.3%) had distractors that works very good, 10 items (28.6%) had good functioning distractors, 5 items (14.3%) had distractors that works fair, 11 items (31.4%) had less good distractors, and four items (11.4%) had not good distractors;

3. Yu-mien Shih's Research

The third research was done by Yu-mien Shih (2010) "An Item Analysis of an English Achievement Test Taken by EFL College Students in

Taiwan" Yu-mien Shih did a research about level of difficulty, index of discrimation and the fuction of distractor on english achievement test taken by EFL college students in Taiwan. The researcher took 32 native Chinese freshmen. The result shown that vocabulary test are the most difficult part whole test and exhibit the most indexful discrimination. A higher presentage of items in listening, reading and grammar tests show little or no discrimination index. More nonsense distractors appear in the listening tests. The outcome of the analysis suggests that the test needs to be revised.

4. Ahmad Hanafi and Suhardi's Journal

Other journal was arranged by Ahmad Hanafi and Suhardi entitled "The Characteristics of the English Test Items in Ne for VHS's in Tabalong Regency South Kalimantan in the Academic Year of 2010/2011". This journal showed that, of the 50 test items, each item contains thematerial, construction, and language aspects which are in general relevant. The weakness in the construction aspect is that pictures, graphs, tables, diagrams and the like do not function well. The results of the quantitative analysis using CTT reveal that the items are good based on the difficulty index, discrimination index, distractor effectiveness, and reliability. The quantitative analysis using IRT shows that 27 items fit IRT with the 3PL model, because the information is more complete and the standard error is smaller than the 1PL and 2 PL models although the sample is not adequate enough. The results of the mapping of the quality of English learning based on the results of the English test in the NE show that out of the six 10 VHSs in Tabalong Regency, SMK Tabalong is at the top. The ranks below are occupied by SMKN 1 Tanjung, SMKN 1 Muara Uya, SMKN 1 Haruai, SMKN 1 Banua Lawas, and SMK An Noor Paliat.

5. The differences and similarities

Based on previous research above, the researcher proposed current research that different from the previous researches. **The difference** the current research with the first previous research is in the object of the research and the variable of the research. Arin Eka Primadani's Research

took an object research at the 7 grade accelerated class at SMPN 3 Malang, while the current research took an object research at MTs N Klaten. And the variable, Arin Eka Primadani's Research did an analysis on the content validity, test reliability, item discrimination, item validity, and the effectiveness of the distracters, while the current research did an analysis on the level of difficulty, discrimination power and the effectiveness of distractor. The similarity between both researches is the research methodology both of them use descriptive quantitative research.

The differences between the current research and the second previous research are in the object and variable of the research. Dian Nurlianto's Research took an object research at the Grade XII IPS SMA Negeri Banyumas, while the current research took an object research at MTs N Klaten. And the variable, Dian Nurlianto's Research did an analysis on validity, reliability, level of difficulty, distinguishing index, and the pattern of answer distributions of the Odd Semester Final Examination Questions in Accounting Economics Subject, while the current research did an analysis on the level of difficulty, discrimination power and the effectiveness of distractor of English Final Test. The similarity between both researches is the research methodology both of them use descriptive quantitative

The differences between the current research and the third previous research are object of the research. The object of current research is EFL College Students in Taiwan, while the previous research is MTs N Klaten. The similarity between both researches is did an analysis on the level of difficulty, index of discrimation and the fuction of distractor .

The differences between the current research and the fourth previous research are the journal apply CTT to calculate quantitative aspect but in the research uses manual computation by using Arikunto's formula. The similarity of both are analyze from quantitative aspects covers level of difficulty, discriminating power and the effectiveness of distractors.

For clear, the similarities and differences about the previous research and this research can be seen on the table 2.1

Table 2.1 The similarities and differences about the previous research and this research

No	Title	Differences	Similarities
1	An Analysis of a Midterm English Test of the Seventh Grade Accelerated Class at SMPN 3 Malang	the object of the research and the variable of the research	the research methodology both of them use descriptive quantitative research
2	Test Item Analysis of the Final Examination on Economics Subject in Grade XII IPS SMA Negeri Banyumas Academic year 2014/2015.	the object and variable of the research	the research methodology both of them use descriptive quantitative
3	An Item Analysis of an English Achievement Test Taken by EFL College Students in Taiwan	object of the research	
4	The Characteristics of the English Test Items in Ne for VHS's in Tabalong Regency South Kalimantan in the Academic Year of 2010/2011	The journal apply CTT to calculate quantitative aspect but in the research uses manual computation by using Arikunto's formula.	Analyze from quantitative aspects covers level of difficulty, discriminating power and the effectiveness of distractors.

CHAPTER III

RESEARCH METHODOLOGY

F. Research Design

This study is shape into a descriptive quantitative study. Sukamolson (2000:2) states quantitative research is a type of research that is explaining phenomena by collecting numerical data that are analyzed using mathematically based methods. Quantitative research is the research with statistical prosedure to reveal the unknown phenomena. This study is descriptive because it describes the difficulty level, discrimitating power and distraction evaluation of the english final test for second grade students of MTs N Klaten. Moreover, this study is quantitative because the researcher used some numerical data which is analyzed statistically. So, this study describe the result of numerical computation of the data.

G. Place and Time of the Research

1. Place of the Research

The research was conducted in MTs N Klaten whic is located in Jl. Ki Ageng Gribig, Gergunung, Klaten Utara.

2. Time of the Research

The research was carried out from April 2017 until February 2018.

H. Subject of the Research

To simplify and reduce the complexity that might occur during the research, the researcher needed to specify the population, sample and the technique of sampling.

1. Population

Population is all respondents of the research subject (Arikunto: 2010). Sangadji and Sopiah (2010: 185) state that population is generalization region consisting of the subject or object with specific qualities and characteristics defined by the researcher to study and then

drawn the conclusions. Population is all individuals which become the source of the sampling.

In this study, the population was all the multiple-choice test items and essay used in the English Final test of the second grade students of MTs N Klaten in the academic year 2016/2017. There are eight classes consist of VIII A, VIII B, VIII C, VIII D, VIII E, VIII F, VIII G and VIII H. Each class has 30-35 numbers of students.

2. Sample

Cresswell (2012: 142) defines a sample as a subgroup of the target population that the researcher plans to study for the generatizing about the target population. A sample was used as an adequate representative of the entire population for the purpose of the study. Arikunto (2005:134) stated if the total number of population is less than 100, it is better for the researcher to take the whole number of population in his research. While the member of population is more than 100, the sample can be taken around 10%-15%, 20%-25% or more from the total number of the population. Based on the explanation above, the researcher take 100 students' test results of the second grade students of MTs N Klaten from 300 members of population.

3. Technique of sampling

Cresswell (2012: 142) state that the strict form of sampling is selecting the individuals from the population who are representative of that population. This research applied random sampling. The procedures to do stratified random sampling according to Rahmatina (2010:79) as follows:

a. Population devided into smaller population which is called stratum.
b. Stratum devided until getting homogenous stratum.
c. Each stratum taken randomly to represent each homogenous stratum.
d. Over all estimation obtained by combination.

To do the sampling, the researcher took 27 students (27%) which are obtained from the lower group, 46 students (46%) which are

obtained from middle group and 27 students (27%) which are obtained from the upper group. So, the research determined 100 students' test result from the calculation above.

The procedure of getting the sample were:

a. From the eight classes of the 8th grade i.e. VIII A, VIII B, VIII C, VIII D, VIII E, VIII F, VIII G and VIII H, the researcher took five class randomly i.e. VIII A, VIII C, VIII E, VIII G and VIII H.

b. From that selected classes, the researcher took 20 students randomly for each class.

I. Technique of Collecting Data

The technique that used in this study was document analysis. Cresswell (2012: 154) considers documents as the represent accurate data consisting of numeric, individual data available in publik records. Those types of data include grade report, school attendance records, students demographic data and census information.

The document that used in this study was the files or data of related information including the English final test items and the students answer sheet. In collecting the data, the researcher did an observation by visiting the school to ask for the students answer sheet and test paper of English final test for the second grade students of MTs N Klaten in the academic year 2016/2017 to be analyzed.

J. Technique of Analizing the Data

The purpose of this research is to describe the quality of the English final test for second grade students of MTs N Klaten by analyzing the students' result of the test. By using certain formulas, the researcher calculated the students' score and transform them into index for each category (difficulty level, discrimitating index and distraction evaluation). By using criteria for each category (difficulty level, discrimitating index and distraction evaluation) the indexes will show the quality of the test. The following was the formulas of item analysis suggested by Arifin and used in the analysis.

1. Difficulty Level

In the multiple choice item, the level of difficulty was calculated by using the following formula:

$$P = \frac{(100)(\sum B)}{N}$$

Where:

P = the index of difficulty (the precentage of students answering the item correctly)

$\sum B$ = the number of students who answer correctly

N = the total of students taking the test

Whereas, the technique to find out the level of difficulty in essay, the researcher used the following formula:

First, determine the mean score of the essay test item by using thefoolowing formula

$$\text{mean} = \frac{\text{the total of students' score for each item}}{\text{Number of students}}$$

The formula of difficulty level as follow:

$$\text{difficulty index} = \frac{\text{mean}}{\text{maximum score of each item}}$$

Difficulty level of an item was interpreted through the following range:

Table 3.3 Index of difficulty level

P	Interpretation
Less than 0,30	Very difficult
0,30 – 0,70	Satisfactory
More than 0,70	Very easy

2. Discrimitating Index

After selecting the upper and lower group, the discriminating index can be analyzed by using this foolowing formula:

$$DP = \frac{WL}{n} - \frac{WH}{n}$$

In which:

DP   = Discrimitating Index

WL   = number of person in lower group who answered the item incorrectly

WH   = number of person in upper group who answered the item incorrectly

n   = number of students in each group

The above formula was used to analyzed the discriminating index of multiple choice item. Whereas, the essay iten can be analyzed by using following formula:

First, the researcher find out the mean score of essay item in each group (lower and upper group), after that calculating the discriminating index by using following formula:

$$DP = \frac{X\ KA - X\ KB}{score\ max}$$

In which:

DP   = discriminating index

X KA   = mean of upper group

X KB   = mean of lower group

Score max = maximum score

Finally, the discriminating index of an item can be interpreted as follow:

Table 3.4. Discriminating index

DP	Interpretation
≥ 0.40	Very good item
$0.30 - 0.39$	Reasonably good item
$0.20 - 0.29$	Marginal item
≤ 0.19	Poor item, to be rejected or improved

3. Distraction Evaluation

Arifin (2012: 282) states that the distractor effective when:

1). The total of students in lower and upper group who choice the distractor not less than:

$$25\% \times \frac{1}{2\sum d} \; x \; (ka + kb)$$

In which:

d   = number of distraction

ka   = upper group students

kb   = lower group students

2). The distracters are chosen by the students in lower group instead of the students in upper group

3). Omit is not more than 10 % of the students.

So, the distracter is effective when it is chosen by ≥ 2 students and omit is not more than 5 studentss. It is also effective when the total number of students in upper group who choose the distracter is less than the total number of students in lower group.the effectiveness of distracters are only applied in analyzing the multiple choice items.

CHAPTER IV

DATA DESCRIPTION, RESEARCH FINDING AND DISCUSSION

A. Data Description

In this chapter, the researcher had analyzed the data gathered from the research. The data was obtained from: 1.) The students' answer sheet and English final test question paper for the second grade students of MTs N Klaten in the academic year 2016/2017. The writer used English final test which was conducted on Tuesday, June 6th 2017, from 07.30 to 09.30 a.m. The test consists of 45 items, 40 multiple choice items and 5 essay items. Having gained the whole needed data, the researcher then did analysis which refers to the statistical data analysis to find out whether or not the test is categorized as a good test.

The good test should good at it item analysis to be a good test. The item analysis contains three components, they are difficulty level, discriminating power and the effectiveness of the distractor. According to Purwanto, good items test is test which have moderate difficulty level, high discriminating power and distractor analysis which work effectively. According to Arikunto (2005: 207), the good items should fulfill of satisfactory difficulty level, satisfactory or good discriminating index and effective distractors (every distractor should be choosen by more 5% of studens' total number).

The first thing to do before conducting analysis was classifying the student based on the test result. The researcher divided 100 students into three classification they are upper group, middle group and lower group. Here, only two groups that is used to analyze, they are upper and lower group. Meanwhile, the middle group is disregarded. The table below is the students position in the group based on the test result.

Table 4.0 The group position based on the test result

Numb	Name	Score	Category	Numb	Name	Score	Category
1	CS31	80	**UPPER**	51	CS 3	48	**MIDDLE**
2	CS 148	78	**UPPER**	52	CS 9	48	**MIDDLE**
3	CS 5	76	**UPPER**	53	CS 19	48	**MIDDLE**
4	CS 147	76	**UPPER**	54	CS 42	48	**MIDDLE**
5	CS 18	74	**UPPER**	55	CS 61	48	**MIDDLE**
6	CS 105	73	**UPPER**	56	CS 66	48	**MIDDLE**
7	CS 106	72,5	**UPPER**	57	CS 82	48	**MIDDLE**
8	CS 13	72	**UPPER**	58	CS 117	48	**MIDDLE**
9	CS 10	70	**UPPER**	59	CS 68	47	**MIDDLE**
10	CS 12	70	**UPPER**	60	CS 103	47	**MIDDLE**
11	CS 53	70	**UPPER**	61	CS 135	47	**MIDDLE**
12	CS 153	70	**UPPER**	62	CS 142	47	**MIDDLE**
13	CS 104	68,5	**UPPER**	63	CS 2	46	**MIDDLE**
14	cs 20	68	**UPPER**	64	CS 6	46	**MIDDLE**
15	CS 57	68	**UPPER**	65	CS 7	46	**MIDDLE**
16	CS 149	68	**UPPER**	66	CS 38	46	**MIDDLE**
17	CS 23	62	**UPPER**	67	CS 52	46	**MIDDLE**
18	CS 145	61	**UPPER**	68	CS 59	46	**MIDDLE**
19	CS 11	60	**UPPER**	69	CS 64	45	**MIDDLE**
20	CS 137	60	**UPPER**	70	CS 99	45	**MIDDLE**
21	CS 110	59,5	**UPPER**	71	CS 58	44	**MIDDLE**
22	CS 54	58	**UPPER**	72	CS 75	44	**MIDDLE**
23	CS 60	58	**UPPER**	73	CS 78	43	**MIDDLE**
24	CS 71	58	**UPPER**	74	CS 29	42	**LOWER**
25	CS 27	56	**UPPER**	75	CS 30	42	**LOWER**
26	CS 33	56	**UPPER**	76	CS 47	40	**LOWER**
27	CS 34	56	**UPPER**	77	CS 122	39	**LOWER**
28	CS 25	54	**MIDDLE**	78	CS 4	38	**LOWER**
29	CS 150	54	**MIDDLE**	79	CS 40	38	**LOWER**
30	CS 70	53	**MIDDLE**	80	CS 163	37	**LOWER**
31	CS 73	53	**MIDDLE**	81	CS 84	36	**LOWER**
32	CS 86	53	**MIDDLE**	82	CS 120	36	**LOWER**
33	CS 102	53	**MIDDLE**	83	CS 154	36	**LOWER**

34	CS 63	52	**MIDDLE**	84	CS 155	35	**LOWER**
35	CS 67	52	**MIDDLE**	85	CS 48	34	**LOWER**
36	CS 126	52	**MIDDLE**	86	CS 89	34	**LOWER**
37	CS 130	52	**MIDDLE**	87	CS 96	34	**LOWER**
38	CS 141	52	**MIDDLE**	88	CS 121	34	**LOWER**
39	CS 100	51	**MIDDLE**	89	CS 114	33	**LOWER**
40	CS 131	51	**MIDDLE**	90	CS 116	33	**LOWER**
41	CS 8	50	**MIDDLE**	91	CS 36	32	**LOWER**
42	CS 21	50	**MIDDLE**	92	CS77	32	**LOWER**
43	CS 28	50	**MIDDLE**	93	CS 133	32	**LOWER**
44	CS 32	50	**MIDDLE**	94	CS 162	32	**LOWER**
45	CS 45	50	**MIDDLE**	95	CS 119	31	**LOWER**
46	CS 108	50	**MIDDLE**	96	CS 92	30	**LOWER**
47	CS 128	50	**MIDDLE**	97	CS 69	29	**LOWER**
48	CS 85	49	**MIDDLE**	98	CS 79	28	**LOWER**
49	CS 101	49	**MIDDLE**	99	CS 125	26	**LOWER**
50	CS 107	49	**MIDDLE**	100	CS 161	25	**LOWER**

B. Research Finding

In this sub-chapter, the researcher presents the findings of the research that aim to answer the research question of this researchabout the difficulty level, discriminating index and distraction evaluation on English Final Test for the second grade students of MTs N Klaten in the academic years 2016/2017 proposed by Arikunto's formula. This item analysis took from 100 students' answer sheet of English Final Test for the second grade students of MTs N Klaten in the academic years 2016/2017 which has devided into three groups, 27 students' answer sheets (27%) from upper group, 46 students' answer sheets (46%) from middle group and 27 students' answer sheets (27%) from lower group. Upper and lower groups are used in the computation, while the middle group is discarded.

a. Difficulty Level

The researcher measures the difficulty level that exists in items of the English final test by using Arikunto's formula. (1995: 212), number

which is indicates the items that difficult or easy called *difficulty index*. Number of difficulty index between 0.00 until 1.0. It is shows the standard of test difficulty. Test with the difficulty index 0.0 show that the test is too hard, in opposite index 1.0 show that the test is too easy.

A good item is the comparison between number of students who got right answer and total number of all the students who tried to answer shows about 0.50. The item could be said not easy and not too difficult. Difficulty level index which accepted to state a satisfactory item is about 0.25 until 0.75. From the analysis, the difficulty level of multiple choice items are 10 items (25%) consider become difficult items, 26 items (65%) consider become satisfactory items and 4 items (10%) are easy items. For the detail result of difficulty level of English final test for second grade students of MTs N Klaten in the academic year 2016/2017 can be seen on table 4.1

Table 4.1 The result of difficulty level of multiple choice items

Difficulty Level (P)	Item Number	Total	Precentage (%)
Difficult (Less than 0,30)	11, 12, 14, 16, 20, 27, 31, 32, 34, 38	10	25
Satisfactory	1, 3, 4, 5, 6, 7, 8, 9, 10, 13, 15, 17, 18, 19, 21, 24, 25, 26, 29, 30, 33, 35, 36, 37, 39, 40	26	65
(0,30 - 0,70)			
Easy (More than 0,70)	2, 22, 23, 28	4	10

While the essay items found that there are 5 items (100%) consider become satisfactory items. For the detail result of difficulty level of English final test for second grade students of MTs N Klaten in the academic year 2016/2017 can be seen on table 4.2

Table 4.2 The result of difficulty level of essay items

Difficulty Level (P)	Item Number	Total	Precentage (%)
Difficult (Less than 0,30)			0
Satisfactory (0,30 - 0,70)	41, 42, 43, 44, 45	5	100
Easy (More than 0,70)			0

b. Discriminating Index

According to Arikunto, discriminating Index is devided into four category; poor, satisfactory, good and excellent. The result of the multiple choice item found that the discriminating index is about 0,00 until 0,20 and have 13 items (32.5%) of negative value. It is consider into poor discriminating index. The discriminating index is about 0,20 until 0,40 have 11 items (27.5%) of negative value. It is consider into satisfactory discriminating index. There are 16 items (40%) have 0,40 until 0,70 discriminating index and it is consider into good discriminating index. The detail result of the discriminating index can be seen on table 4.3

Table 4.3 The result of discriminating index of multiple choice items

Discriminating Index (D)	Item Number	Total	Precentage (%)
Poor (0,00 - 0,20)	12, 13, 14, 16, 20, 22, 23, 24, 27, 31, 32, 34, 38	13	32,5
Satisfactory (0,20 - 0,40)	7, 1, 2, 28, 17, 18, 26, 11, 19, 21, 40,	11	27,5
Good (0,40 - 0,70)	3, 4, 5, 6, 8, 9, 10, 15, 25, 29, 30, 33, 35, 36, 37, 39	16	40
Excellent (more than 0,70)	-	0	0

While the essay items, there are 2 item (40%) considered as poor item and 3 items (60%) are considered as satisfactory discriminating index. The detail result of the discriminating index can be seen on table 4.4

Table 4.4 The result of discriminating index of essay items

Discriminating Power (D)	Item Number	Total	Precentage (%)
Poor (0,00 - 0,20)	41, 42	2	40
Satisfactory (0,20 - 0,40)	43, 44, 45	3	60
Good (0,40 - 0,70)		0	0
Excellent (more than 0,70)		0	0

c. Distraction Evaluation

English final test items for second grade students of MTs N Klzten in the academic year 2016/2017 covers 40 multiple choice items with four possible answers. There are one correct answer and three distractors for every item. The total number of distractor in the English final test items for second grade students of MTs N Klzten in the academic year 2016/2017 is 120 distractors. Based on Arikunto, distractor could be considered as a good distractor if it attracts more than 5% (2 students) of total students who tried to answer. A good distractor should be chosen by more students in lower group than students in upper group.

From the total number of distractor, there are 86 distractors (71,67) which is categorized as effective distractors, 7 distractors (5,83) which is categorized as less effective distractors and 27 distractors (22,50) categorized as ineffective distractors. For the detail result of the effectiveness of distractor can be seen on table 4.5.

Table 4.5 Result of the effectiveness of distractor of multiple choice
items

Item	Correct	Effectiveness of Distractor		
Number	Answer	Effective	Less Effective	Ineffective
1	D	A B C		
2	C	A B		D
3	D	A B C		
4	D	A B C		
5	A	B C D		
6	B	A C D		
7	C	A B		D
8	D	A B		C
9	A	B C D		
10	A	C	B D	
11	B	A D		C
12	C	D B		A
13	D	B	C	A
14	D	C		B D
15	B	C D		A
16	C	B		A D
17	B	C D		A
18	A	B C D		
19	D	A B	C	
20	A	B		C D
21	D	A B C		
22	B	A	D	C
23	B	C		A D
24	D	C	A	B
25	C	A B D		
26	A	B D		C
27	B	D		A C
28	C	A B		D
29	C	A B D		

30	B	A C D		
31	C	B D		A
32	A	B D	C	
33	C	A B		D
34	D	B		A C
35	B	A C D		
36	B	A C D		
37	A	B C D		
38	B	C D		A
39	B	A C D		
40	A	B C D		
Total Option		86	7	27
(Precentage)		(71,67)	(5,83)	(22,50)

The result of the analysis based on difficulty level on multiple choice items show that 14 items (28%) are considered as easy item, 10 items (25%) consider become difficult items, 26 items (65%) consider become satisfactory items and 4 items (10%) are easy items. While the essay items found that there are 5 items (100%) consider become satisfactory items. The result of discriminating index on multiple choice items found that 13 items (32.5%) have poor discriminating index, 11 items (27.5%) have satisfactory discriminating index, 16 items (40%) have good discriminating index and no item has excellent discriminating index. While the essay items, there are 2 item (40%) considered as poor item and 3 items (60%) are considered as satisfactory discriminating index. The result of effectiveness distractor is 86 distractors (71,67) which is categorized as effective distractors, 7 distractors (5,83) which is categorized as less effective distractors and 27 distractors (22,50) categorized as ineffective distractors.

From the data above, the items can be devided into three criterias as follows accepted without revision, revised and refused. The item which is accepted without revision if it has satisfactory difficulty level, satisfactory or good discriminating index, and effective distractor. The

item which is revised because it only fulfills two of three criterias. The item which is refused because it has easy or difficult of difficulty level, poor discriminating index and ineffective distractor. According to the data, for the multiple choice items, there is 14 items (40%) included accepted item, 15 item (37.5%) included revised item and 11 items (27.5) considered refused item. The result of grouping in the multiple choice items can be seen on the table 4.6

Table 4.6 The grouping of multiple choice items analysis

Criteria of Item	Item Number	Total	Percentage (%)
Good (Accepted)	1, 3, 4, 5, 6, 18, 21, 25, 30, 35, 36, 37, 39, 40	14	35
Medium (Revised)	7, 8, 9, 10, 11, 13, 15, 17, 19, 24, 26, 28, 29, 32, 33	15	37,5
Poor (Refused)	2, 12, 14, 16, 20, 22, 23, 27, 31, 34, 38	11	27,5

While the essay items, there is 3 items (60%) included accepted item, 2 item (40%) included revised item. The result of grouping in the essay items can be seen on the table 4.7

Table 4.7 The grouping of essay items analysis

Criteria of Item	Item Number	Total	Percentage (%)
Good (Accepted)	43, 44, 45	3	60
Medium (Revised)	41, 42	2	40

Poor (Refused)		0	0

C. Discussion

The researcher discusses the research findings with the other relevant references. In order to justify the research findings, the researcher discusses them with reference to theory related to the study of multiple choice item toward English final test items for second grade students of MTs N Klaten in academic year 2016/2017.

1. The Level of Difficulty

The percentages of the level of difficulty on multiple choice item from table 4.1 were 10% easy items, 65% satisfactory items and 25% difficult items. While The percentages of the level of difficulty on essay items from table 4.2 were 100% satisfactory items. Item must be in appropriate difficulty for the students to whom it is administered. If possible, items should have indices of difficulty no less than 0.30 and no greater than 0.70. It is desirable to have most items in the 0.30-0.70 range of difficulty. Too difficult or too easy items contribute little to the discriminatingindexof a test. The level of difficulty for each item has the relationship and effect in arranging the test items. From table 6.3, it can be seen thatEnglishfinal test for Eighth grade students of MTs N Klaten has bad arrangement of difficulty level test. The test is started from satisfactory question then followed by difficult question in number 11. Moreover, the easy item test appeared in the middle item number.

Djiwandono (2008:220) states that giving the difficult question which makes the students think harder and consumes the more time to answer will lead to have bad effect, because they will feel inferior and afraid while doing the difficult items in the test and it also affect to the next questions. The difficult test items must be arranged in the last item so that the students feel confidence to answer because of having done theprevious questions. Moreover, if the students feel troubled to answer the last items, it will not affect to the previous items.

2. The DiscriminationIndex

From the table 4.3 , it can be found that the multiple choice items for 1, 2, 7, 11, 17, 18, 19, 21, 26, 28, 40 included as the satisfactory discrimination, number 3, 4, 5, 6, 8, 9, 10, 15, 25, 29, 30, 33, 35, 36, 37 and 39 included as the good discrimination. While the essay items, for 43, 44 and 45 included as satisfactory discrimination. Because they had the information about the differences in the performance of the students. The teacher or the test maker can keep saving those items to give in the next test because they had the information about the differences in the performance of the students.. In other hand, the multiple choice items numbers 12, 13, 14, 16, 20, 22, 23, 24, 27, 31, 32, 34, and 38 are considered poor discriminability. While the essay items, for number 41 and 42 also considered into poor discrimination. Because, the items cannot give the information about the differences of the performance among the students.

On contrary, the students who are supposed to have high ability got the wrong item. Besides, the students who are in lower ability got the correct item. The discrimination is important feature of test. It is the capacity to discriminate among different candidates, reflect the differences in the performance of the individuals in the group and distinguish among the students who are in high ability or got the item correct and those who are in lower ability to respond the items correctly. The higher discrimination index of test item, the better it is. Sudjiono (1996:408) state that following up must be done by the teacher or the test maker after analyzing the discriminationindexfor eachitem

The items which have good discrimination index (satisfactory, good and excellent classification) should be kept in item test bank, so that can be used later. The items which are categorized as the poor discrimination should have been revised then used later or dropped. The very poor discrimination of test must be dropped or not to be used later

3. **The effectiveness of Distractor for Each Items**

   Commonly, the multiple-choice question has the basic structure; a stem, option which consists of the answer and distractor. All of the incorrect options, or distractors, should actually be distracting. Preferably, each distractor should be selected by a greater proportion of the lower group than that of the upper group. From table 4.5, the distractor A shows in item numbers 1, 2, 3, 4, 6, 7, 8, 11, 19, 21, 22, 25, 28, 29, 30, 33, 35, 36 and 39. Distractor B shows in item numbers 1, 2, 3, 4, 5, 7, 8, 9, 12, 13, 16, 18, 19, 20, 21, 25, 26, 28, 29, 31, 32, 33, 34, 37 and 40. Distractor C shows in item numbers 1, 3, 4, 5, 6, 9, 10, 14, 15, 17, 18, 21, 23, 24, 30, 35, 36, 37, 38, 39 and 40. Distractor D shows in item numbers 5, 6, 9, 11, 12, 15, 17, 18, 25, 26, 27, 29, 30, 31, 32, 35, 36, 37, 38, 39 and 40. They are categorized as the effective distractor, because there are more than 5% students who choose those distractors. Besides that, distractor A shows in item numbers 24. Distractor B shows in item numbers 10. Distractor C shows in item numbers 13, 19 and 32. D in item numbers 10 and 22 . They are called as ineffective distactors, because those distractors are chosen by less than 5% from all the students. The other distractors which have not been mentioned above are called as omit because no students are interested in choosing. They should be deleted or revised

   Sudjiono (1996:409) states that the distractor functions well while it is chosen by at least 5% from the all students. If a distractorelicits very few or no responses, then it may not be functioning as a distractor and should be replaced with a more attractive option. In addition, some distractor may be too appealing and causing the items to be too difficult. Very often items which have been rejected as having inappropriate difficulty, discriminability or variability can be redeemed by the revision of one or two response options. Sudjiono (1996:410-413) states that if there is no one chooses the provided distractor, it means that the distractor cannot functioned well. It must be dropped or revised. Haladyna (2004:99) suggests how to write good options (either distractor or key answer) as follows:

a. Develop as many effective options as the test maker can, but two or three may be sufficient

b. Vary the location of the right answer according to the number of options. Assign the position of the right answer randomly

c. Place option independent; choices should not be overlapping

d. Keep the options homogeneous in content and grammatical structure

e. Keep the length of options about the same

f. Make distracters plausible

g. Use typical errors of the students when writing distractor

h. Avoid option that give clues to the right answer

According to Arikunto (2005: 207), the good items should fulfill of satisfactory difficulty level (P = 0.30-0.70), satisfactory or good discriminating index (D = 0.20-0.70) and effective distractors (every distractor should be choosen by more 5% of studens' total number). The item that included in the category of good items are accepted and it can be used without revising. The result of the analysis shows that there are 35% of multiple choice items and 60% of essay items that included in the category of good item. The item that included in the category of good items are item number 1, 3, 4, 5, 6, 18, 21, 25, 30, 35, 36, 37, 39, 40, 43, 44 and 45. The medium items are the item which fulfill of two of the three criterias as a good item. There are 37.5% of multiple choice items and 40% of essay items included as medium items. The item that included in the category of medium items are item number 2, 7, 8, 9, 10, 11, 13, 15, 17, 19, 24, 26, 28, 29, 32, 33, 41 and 42. The items that included in the category of medium item can be properly used after revised. The poor items are the item which fulfill of one or nor of the three criterias as a good item. There are 27.5% of multiple choice items are included as poor items. The item that included in the category of poor items are item number 2, 12, 14, 16, 20, 22, 23, 27, 31, 34 and 38. The items that included in the category of poor item should be refused or it can be used after got revise.

More explanation can be read as follows:

Discussion item number 1

1. Moon Hotel has ... rooms than Sun Hotel.

    a. Few   c. Much

b. Fewer   d. More

No Item	Opt	UPPER	LOWER
1	A	0	2
	B	3	4
	C	4	9
	D*	20	12
Total		27	27

1) The difficulty level (P) is 0.59. It means that the item is categorize into satisfactory difficulty level since only 32 students who get the right answer.

2) The discriminating index (D) is 0.20. It discriminates upper and lower group students in satisfactory level.

3) All of distractors are functioning well since these attract more students in lower group. The distractor analysis as follow:

    a) Distractor A = acceptable

    b) Distractor B = acceptable

    c) Distractor C = acceptable

In short, this item is a good item because it can fulfill all the aspects to be a good item which is used in the test. It has satisfactory level of difficulty, satisfactory discriminating index, and all of the distracters are functioning effectively.

Discussion item number 2

2. The room rate of Sun Hotelis... than of Moon Hotel

a. More expensive   c. Cheaper

b. Expensive   d. Cheap

No Item	Opt	UPPER	LOWER
	A	1	3
	B	0	7
2	C*	26	16
	D	0	1
Total		27	27

1) The difficulty level (P) is 0.74. It means that the item is categorize into easy difficulty level since only 42 students who get the right answer.

2) The discriminating index (D) is 0.44. It discriminates upper and lower group students in good level.

3) Distractor A and B are effective. They attract more students in lower group. While distractor D is functioning poorly since it only attracts one students of the total number of studentsThe distractor analysis as follow:

a) Distractor A = acceptable

b) Distractor B = acceptable

c) Distractor D= Revised

This item is categorized as poor item because it is an easy item although it has satisfactory level of discriminating index and two distractors are functioning well.

Discussion item number 3

3. The facilities of Moon Hotel are ... than Sun Hotel

a. Fewest   c. Much

b. Fewer   d. More

No Item	Opt	UPPER	LOWER
3	A	2	9
	B	1	7
	C	6	6
	D*	19	5
Total		27	27

1) The difficulty level (P) is 0.44. It means that the item is categorize into satisfactory difficulty level since only 24 students who get the right answer.

2) The discriminating index (D) is 0. 52. It discriminates upper and lower group students in good level.

3) Distractor A and B are effective. They attract more students in lower group. While distractor C is is not functioning as intended because it attracts some students from lower and upper group in equal number. The distractor analysis as follow:

a) Distractor A = acceptable

b) Distractor B = acceptable

c) Distractor C = Revised

By considering the difficulty level, discriminating index and the effectiveness of distractors, this item has become a good item.

Discussion item number 4

4. Sun Hotel has... room types than Moon Hotel

a. Much   c. More

b. Little   d. Fewer

No Item	Opt	UPPER	LOWER

4	A	2	4
	B	4	9
	C	4	10
	D*	17	3
Total		27	27

1) The difficulty level (P) is 0.37. It means that the item is categorize into satisfactory difficulty level since only 20 students who get the right answer.

2) The discriminating index (D) is 0. 48. It discriminates upper and lower group students in good level.

3) All of the distractors are fungtioning well since these attract more students in lower group. The distractor analysis as follow:

a) Distractor A = acceptable

b) Distractor B = acceptable

c) Distractor C = acceptable

By considering the difficulty level, discriminating index and the effectiveness of distractors, this item has become a good item. It has satisfactory difficulty level, has good level of discriminating index and all of the distractors are fungtioning effectivelly.

Discussion item number 5

5. The writer writes the text in order to...

a. Describe her friend's style and her new shoes

b. Explain international shoes

c. Share her experience with her friend

d. Tell Blowfish shoes product

No Item	Opt	UPPER	LOWER
5	**A***	24	11
	B	0	2

	C	1	7
	D	2	7
Total		27	27

1) The difficulty level (P) is 0.65. It means that the item is categorize into satisfactory difficulty level since only 35 students who get the right answer.

2) The discriminating index (D) is 0. 48. It discriminates upper and lower group students in good level.

3) All of the distractors are fungtioning well since these attract more students in lower group. The distractor analysis as follow:

a) Distractor B = acceptable

b) Distractor C = acceptable

c) Distractor D = acceptable

It has satisfactory difficulty level, has good level of discriminating index and all of the distractors are fungtioning effectivelly. By considering the difficulty level, discriminating index and the effectiveness of distractors, this item has become a good item.

Discussion item number 6

6. The writer's friend has just bought...from Blowfish shoes product

   a. A new match shoes   c. A trendy and attractive shoes

   b. A new stylish foot legs   d. A brand and bright color shoes

No Item	Opt	UPPER	LOWER
6	A	0	6
	B*	27	10
	C	0	3
	D	0	8
Total		27	27

1) The difficulty level (P) is 0.67. It means that the item is categorize into satisfactory difficulty level since only 37 students who get the right answer.

2) The discriminating index (D) is 0. 63. It discriminates upper and lower group students in good level.

3) All of the distractors are fungtioning well since these attract more students in lower group. The distractor analysis as follow:

a) Distractor A = acceptable

b) Distractor C = acceptable

c) Distractor D = acceptable

It has satisfactory difficulty level, has good level of discriminating index and all of the distractors are fungtioning effectivelly. So, this item has become a good item.

Discussion item number 7

7. Why does the writer admire her friend?

a. She likes wearing internationally recognized shoes

b. She always wants to be a trendy and attractive woman

c. She has the most suitable shoes on her physical appearance

d. She really has a perfect appearance with her wonderfull shoes

No Item	Opt	UPPER	LOWER
7	A	1	7
	B	2	8
	C*	12	5
	D	12	7
Total		27	27

1) The difficulty level (P) is 0.31. It means that the item is categorize into satisfactory difficulty level since only 17 students who get the right answer.

2) The discriminating index (D) is 0. 26. It discriminates upper and lower group students in satisfactory level.

3) Distractor A and B are effective. They attract more students in lower group. While distractor D is is not functioning as intended because it attracts more students from upper group. The distractor analysis as follow:

a) Distractor A = acceptable

b) Distractor B = acceptable

c) Distractor D = revised

This item has become a good item. It has satisfactory difficulty level and has satisfactory level of discriminating index, although only two distractors that fungtioning effectivelly.

Discussion item number 8

8. What made the students sad?

   a. The strong sun   c. The distance of the beach

   b. The heavy garbage   d. The condition of the beach

No Item	Opt	UPPER	LOWER
8	A	1	4
	B	1	9
	C	6	5
	D*	19	8
Total		27	27

1) The difficulty level (P) is 0.50. It means that the item is categorize into satisfactory difficulty level since only 27 students who get the right answer.

2) The discriminating index (D) is 0.41. It discriminates upper and lower group students in good level.

3) All of the distractors are fungtioning well since these attract more students in lower group. The distractor analysis as follow:

a) Distractor A = acceptable

b) Distractor B = acceptable

c) Distractor C = acceptable

This item is a good item because it can fulfill all the aspects to be a good item which is used in the test. It has satisfactory level of difficulty, good discriminating index, and all of the distracters are functioning effectively.

Discussion item number 9

9. Mrs. Sumiyati is probably a...

| a. Teacher | c. Bus driver |
| b. Students | d. Beach owner |

No Item	Opt	UPPER	LOWER
9	A*	20	5
	B	1	3
	C	5	15
	D	1	4
Total		27	27

1) The difficulty level (P) is 0.46. It means that the item is categorize into satisfactory difficulty level since only 25 students who get the right answer.

2) The discriminating index (D) is 0.56. It discriminates upper and lower group students in good level.

3) All of the distractors are fungtioning well since these attract more students in lower group. The distractor analysis as follow:

a) Distractor B = acceptable

b) Distractor C = acceptable

c) Distractor D = acceptable

This item is a good item because it can fulfill all the aspects to be a good item which is used in the test. It has satisfactory level of difficulty, good discriminating index, and all of the distracters are functioning effectively.

Discussion item number 10

10. What is the purpose of the text?

a. To tell about the writer's experience of having a beach clean up

b. To let the reader know about the condition of Maron Beach

c. To explain the readers about the beauty of Maron Beach

d. To give the readers directions to reach the beach

No Item	Opt	UPPER	LOWER
10	A*	14	3
	B	6	6
	C	3	13
	D	4	4
Total		27	27

1) The difficulty level (P) is 0.31. It means that the item is categorize into satisfactory difficulty level since only 17 students who get the right answer.

2) The discriminating index (D) is 0.41. It discriminates upper and lower group students in good level.

3) Distractor C and D are effective. They attract more students in lower group. While distractor B is is not functioning as intended because it attracts some students from lower and upper group in equal number. The distractor analysis as follow:

a) Distractor B = revised

b) Distractor C = acceptable

c) Distractor D = acceptable

By considering the difficulty level, discriminating index and the effectiveness of distractors, this item has become a good item.

Discussion item number 11

11. What is the specific characteristic of the car?

   a. It is completed with AC, TV, DVD and GPS tracking device
   b. It wil stop automatically if the front car stop accidently
   c. The right and theleft lamp were pretty designed
   d. The car is well designed so it seemsvery strong and gentle

No Item	Opt	UPPER	LOWER
11	A	12	21
	B*	10	1
	C	1	0
	D	4	5
Total		27	27

1) The difficulty level (P) is 0.20. It means that the item is categorize into difficult level since only 11 students who get the right answer.

2) The discriminating index (D) is 0.33. It discriminates upper and lower group students in satisfactorylevel.

3) Distractor A and D are effective. They attract more students in lower group. While distractor C is not functioning as intended because it attracts more students from upper group. The distractor analysis as follow:

a) Distractor A = acceptable

b) Distractor C = revised

c) Distractor D = acceptable

In short, this item become poor item because it has satisfactory discriminating index. Meanwhile, the difficulty level is inappropriate and one distractor works ineffectively.

Discussion item number 12

12. What can we say about the writer's car?

a. It is an unlimited edition car   c. It is a high-tech car

b. It is an inexpensive car   d. It is an imported car

No Item	Opt	UPPER	LOWER
	A	9	8
	B	10	12
12	C*	6	2
	D	2	4
Total		27	27

1) The difficulty level (P) is 0.15. It means that the item is categorize into difficult level since only 8 students who get the right answer.

2) The discriminating index (D) is 0.15. It discriminates upper and lower group students in poor level.

3) Distractor Band D are effective. They attract more students in lower group. While distractor A is not functioning as intended because it attracts more students from upper group. The distractor analysis as follow:

a) Distractor A = revised

b) Distractor B = acceptable

c) Distractor D = acceptable

It has appropriate difficulty level and discriminating level. It has two distractor which is functioning as intended and one distractor is not functioning effectively. So, this item is considered as a poor item.

Discussion item number 13

13. "... we will feel comfortable when we <u>ride</u> it". The underline word has the same meaning with ....

a. Trip   c. Vacation

b. Picnic   d. Drive

No Item	Opt	UPPER	LOWER
13	A	4	3
	B	1	6
	C	2	2
	D*	20	16
Total		27	27

1) The difficulty level (P) is 0.67. It means that the item is categorize into satisfactory difficulty level since only 23 students who get the right answer.

2) The discriminating index (D) is 0.15. It discriminates upper and lower group students in poor level.

3) Distractor B is working effective. It is attract more students in lower group. Distractor A and C are not functioning as intended because distractor A attracts more students from upper group while distractor C attracts some students from lower and upper group in equal number. The distractor analysis as follow:

a) Distractor A = revised

b) Distractor B = acceptable

c) Distractor C = revised

Although it has satisfactory difficulty level, it has poor level of discriminating index. It has only one distractor which is functioning as intended and most of the distractor are not functioning effectively. So it can be categorized into poor item.

Discussion item number 14

14. What is the main idea of the text?

    a. Having beautiful hill   c. On the way hill

    b. Preparing a picnic   d. Having a camp

No Item	Opt	UPPER	LOWER
	A	6	5
	B	4	9
14	C	15	10
	D*	2	3
Total		27	27

1) The difficulty level (P) is 0.07. It means that the item is categorize into difficult level since only 4 students who get the right answer.

2) The discriminating index (D) is -0.07. It discriminates upper and lower group students in poor level.

3) Distractor B is working effective. It is attract more students in lower group. Distractor A and C are not functioning as intended because distractor A attracts more students from upper group while distractor C attracts some students from lower and upper group in equal number. The distractor analysis as follow:

a) Distractor A = revised

b) Distractor B = acceptable

c) Distractor C = revised

    By considering the difficulty level, discriminating index and the effectiveness of distractors, this item has become a poor item. It has difficult level, poor discriminating index and only has one distractor which working effectively.

Discussion item number 15

15. " The day before yesterday, I went to Siska's house..." the underline phrase means ...

a. Two days later   c. Next two days

b. Two days ago   d. Last day

No Item	Opt	UPPER	LOWER
15	A	7	6
	B*	16	5
	C	0	3
	D	4	13
Total		27	27

1) The difficulty level (P) is 0.39. It means that the item is categorize into satisfactory difficulty level since only 21 students who get the right answer.

2) The discriminating index (D) is 0.41. It discriminates upper and lower group students in good level.

3) Distractor C and D are working effective. It is attract more students in lower group. Distractor A is not functioning as intended because it attracts more students from upper group. The distractor analysis as follow:

a) Distractor A = revised

b) Distractor C = acceptable

c) Distractor D = acceptable

     By considering the difficulty level, discriminating index and the effectiveness of distractors, this item has become a good item.

Discussion item number 16

16. Which statement is TRUE according to the text?

     a. The writer brought blanket, jacket, towel, jeans and swimsuit

     b. The writer went to Siska's house in the Sunday morning at 7 a.m

     c. Nini was gone so all of her friends looked for her

     d. In the end, Nini could be contacted by thewriter

No Item	Opt	UPPER	LOWER
16	A	12	7
	B	4	11
	C*	8	7
	D	4	2
Total		27	27

1) The difficulty level (P) is 0.28. It means that the item is categorize into difficult level since only 15 students who get the right answer.

2) The discriminating index (D) is 0.04. It discriminates upper and lower group students in poor level.

3) Distractor A and D are not working effective. It is attract more students in upper group. Distractor B is functioning as intended because it attracts more students from lower group. The distractor analysis as follow:

a) Distractor A = revised

b) Distractor B = acceptable

c) Distractor D = revised

It has difficult level and it has poor level of discriminating index. It has only one distractor which is functioning as intended and most of the distractor are not functioning effectively. So it can be categorized into poor item.

Discussion item number 17

17. Why did Nini go home early?

   a.   Because she did not bring a mobile phone

   b.   Because everyone did not care about her

   c.   Because her father was sick

   d.   Because Lili could contact her

No Item	Opt	UPPER	LOWER
	A	5	1
17	B*	15	6
	C	0	5
	D	7	15
Total		27	27

1) The difficulty level (P) is 0.39. It means that the item is categorize into Satisfactory difficulty level since only 21 students who get the right answer.

2) The discriminating index (D) is 0.33. It discriminates upper and lower group students in Satisfactory level.

3) Distractor A is not working effective. It is attract more students in upper group. Distractor C and D are functioning as intended because it attracts more students from lower group. The distractor analysis as follow:

a) Distractor A = revised

b) Distractor C = acceptable

c) Distractor D = acceptable

By considering the difficulty level, discriminating index and the effectiveness of distractors, this item has become a good item.

Discussion item number 18

18. How long was the trip from Siska's home to the hill?

a. Half an hour   c. Two hours

b. One hour   d. Three hours

No Item	Opt	UPPER	LOWER

18	A*	16	6
	B	7	13
	C	3	6
	D	1	2
Total		27	27

1) The difficulty level (P) is 0.39. It means that the item is categorize into Satisfactory difficulty level since only 21 students who get the right answer.

2) The discriminating index (D) is 0.33. It discriminates upper and lower group students in Satisfactory level.

3) Distractor A is not working effective. It is attract more students in upper group. Distractor C and D are functioning as intended because it attracts more students from lower group. The distractor analysis as follow:

a) Distractor A = revised

b) Distractor C = acceptable

c) Distractor D = acceptable

This item has become a good item. It has satisfactory difficulty level and has satisfactory level of discriminating index, although only two distractors that fungtioning effectivelly.

Discussion item number 19

19. What is the text tell us about?

a. Message of Nina on school from morning to afternoon

b. Message of Jojo's interest of picking his sister up at 3.30 p.m

c. News of scholl timearound 30 minutes

d. Message to picking Nina' up at 3.30 p.m

No Item	Opt	UPPER	LOWER
19	A	2	3

	B	10	13
	C	2	2
	D*	13	7
Total		27	27

1) The difficulty level (P) is 0.37. It means that the item is categorize into Satisfactory difficulty level since only 20 students who get the right answer.

2) The discriminating index (D) is 0.22. It discriminates upper and lower group students in Satisfactory level.

3) Distractor C is not not functioning as intended because it attracts some students from lower and upper group in equal   number. Distractor A and B are functioning as intended because it attracts more students from lower group. The distractor analysis as follow:

a) Distractor A = acceptable

b) Distractor B = acceptable

c) Distractor C = revised

By considering the difficulty level, discriminating index and the effectiveness of distractors, this item has become a good item. It has satisfactory difficulty level and has satisfactory level of discriminating index, although only two distractors that fungtioning effectivelly.

 Discussion item number 20

20. What time doesNina usually go home from school?

a.   At 3 p.m

b.   At 3.30 p.m

c. At 4 p.m

d. At 4.30 p.m

No Item	Opt	UPPER	LOWER
20	**A***	5	2
	B	14	19
	C	8	6

	D	0	0
Total		27	27

1) The difficulty level (P) is 0.17. It means that the item is categorize into difficult level since only 7 students who get the right answer.

2) The discriminating index (D) is 0.04. It discriminates upper and lower group students in poor level.

3) Distractor B is functioning well. Distractor C is functioning as intended because it attracts more students from lower group. Distractor D is not functioning at all because it attracts no student. The distractor analysis as follow:

a) Distractor B = acceptable

b) Distractor C = revised

c) Distractor D = refused

It can be categorized into poor item. It has difficult level and it has poor level of discriminating index. It has only one distractor which is functioning as intended and most of the distractor are not functioning effectively.

Discussion item number 21

21. How many days are the students asked to bring are ruler?

a. Every day   c. Thirteen days

b. Eleven days   d. Two days

No Item	Opt	UPPER	LOWER
21	A	0	5
	B	1	2
	C	3	7
	D*	23	13
Total		27	27

1) The difficulty level (P) is 0.67. It means that the item is categorize into satisfactory difficulty level since only 36 students who get the right answer.

2) The discriminating index (D) is 0.37. It discriminates upper and lower group students in satisfactory level.

3) All of the distractors are fungtioning well since these attract more students in lower group. The distractor analysis as follow:

a) Distractor A = acceptable

b) Distractor B = acceptable

c) Distractor C = acceptable

From the difficulty level, discriminating indexand distraction evaluation, this item can be categorized into good item.

Discussion item number 22

22. What must the students do if they do not bring are ruler?

   a. They must remind others students

   b. They must buy in the stationary shop of the school

   c. They will be punished by the teacher

   d. They will be fined by the teacher

No Item	Opt	UPPER	LOWER
22	A	2	6
	B*	24	19
	C	0	1
	D	1	1
Total		27	27

1) The difficulty level (P) is 0.80. It means that the item is categorize into easy difficulty level since only 36 students who get the right answer.

2) The discriminating index (D) is 0.19. It discriminates upper and lower group students in poor level.

3) Distractor A is work as intended, while distractor C is not not functioning as intended because it attracts some students from lower and upper group in equal number. distractor D is functioning poorly since it only attracts one students of the total number of students.The distractor analysis as follow:

a) Distractor A = acceptable

b) Distractor C = revised

c) Distractor D = revised

This item is categorized into poor item. It has poor level of discriminating index and easy level of difficulty. While the distractors, there is only one distractor that functioning well.

 Discussion item number 23

23. Where will you fint the text?

a. At bus station   c. At market

b. At school   d. At park

No Item	Opt	UPPER	LOWER
23	A	0	0
	B*	27	24
	C	0	2
	D	0	1
Total		27	27

1) The difficulty level (P) is 0.94. It means that the item is categorize into easy difficulty level since only 51 students who get the right answer.

2) The discriminating index (D) is 0.07. It discriminates upper and lower group students in poor level.

3) Distractor A does not work at all because it attracts no students. Distractor B works as intended, distractor D is functioning poorly since it only attracts one students of the total number of students. This is the analysis of the distractor:

a) Distractor A = refused

b) Distractor B = acceptable

c) Distractor D = revised

In short, this item is poor item. It has appropriate discriminating index, but it has poor level of difficulty. It only has one distarctor which is functioned effectively.

Discussion item number 24

24. What did they feel on their way to find a big city?

a. Happy   c. Sad

b. Glad   d. Tired

No Item	Opt	UPPER	LOWER
24	A	11	11
	B	4	1
	C	0	8
	D*	12	7
Total		27	27

1) The difficulty level (P) is 0.35. It means that the item is categorize into satisfactory difficulty level since only 51 students who get the right answer.

2) The discriminating index (D) is 0.19. It discriminates upper and lower group students in poor level.

3) Distractor A is not not functioning as intended because it attracts some students from lower and upper group in equal number. Distractor B is functioning poorly since it attracts more students in the upper group.

While distractor C is functioning well. This is the analysis of the distractor:

a) Distractor A = revised

b) Distractor B = revised

c) Distractor D = acceptable

This item is considered as poor item. It has appropriate difficulty level but it has poor discriminating index and only has one effective distractor .

Discussion item number 25

25. Why did one of the frogs climb on the other's back?

a.   It feel tired   c. It could not see the city

b.   It was a hot day   d. It thought it was worthwhile

No Item	Opt	UPPER	LOWER
	A	2	6
25	B	3	7
	C*	22	6
	D	0	8
Total		27	27

1) The difficulty level (P) is 0.52. It means that the item is categorize into satisfactory difficulty level since only 28 students who get the right answer.

2) The discriminating index (D) is 0.59. It discriminates upper and lower group students in good level.

3) All of the distractors are fungtioning well since these attract more students in lower group. The distractor analysis as follow:

a) Distractor A = acceptable

b) Distractor B = acceptable

c) Distractor D = acceptable

From the difficulty level, discriminating indexand distraction evaluation, this item can be categorized into good item.

Discussion item number 26

26. "... and at last they <u>set off</u> to the city." The underline phrase means ...

    a. Left   c. Decided

    b. Gave up   d. Stopped

No Item	Opt	UPPER	LOWER
26	A*	15	7
	B	4	10
	C	3	1
	D	5	9
Total		27	27

1) The difficulty level (P) is 0.41. It means that the item is categorize into satisfactory difficulty level since only 28 students who get the right answer.

2) The discriminating index (D) is 0.30. It discriminates upper and lower group students in satisfactory level.

3) Distractors B and D are fungtioning well since these attract more students in lower group. Distractor C is not functioning well because it is attract more students in upper group. The distractor analysis as follow:

a) Distractor B = acceptable

b) Distractor C = revised

c) Distractor D = acceptable

This item has become a good item. It has satisfactory difficulty level and has satisfactory level of discriminating index, although only two distractors that fungtioning effectivelly.

Discussion item number 27

27. What is the moral value of the text?

a. Never do something useless with your friends

b. Never trust within a single opinion without other evidence

c. We have to accept whatever information we receive

d. We canalways ask someone's opinions for anything

No Item	Opt	UPPER	LOWER
27	A	11	8
	B*	3	3
	C	6	5
	D	6	11
Total		27	27

1) The difficulty level (P) is 0.11. It means that the item is categorize into difficult level since only 9 students who get the right answer.

2) The discriminating index (D) is 0.04. It discriminates upper and lower group students in poor level.

3) Distractors A and C are not fungtioning well since these attract more students in upper group. Distractor D is functioning as intended. The distractor analysis as follow:

a) Distractor A = revised

b) Distractor C = revised

c) Distractor D = acceptable

This item is considered as really poor one. It is not only have inappropriate level of difficulty and discriminating index, but two of three distractors work ineffectively.

Discussion item number 28

28. The notice means that ...

a. Visitors must fish far from place

b. The people may only feed the fish

c. Fisitors must not fish in the place

d. Everibody can fish in the place

No Item	Opt	UPPER	LOWER
	A	1	2
28	B	0	8
	C*	26	17
	D	0	0
Total		27	27

1) The difficulty level (P) is 0.80. It means that the item is categorize into easy difficulty level since only 43 students who get the right answer.

2) The discriminating index (D) is 0.33. It discriminates upper and lower group students in satisfactory level.

3) Distractors A and B are fungtioning well since these attract more students in lower group. Distractor D is not functioning at all because it attracts no student. The distractor analysis as follow:

a) Distractor A = acceptable

b) Distractor B = acceptable

c) Distractor D = refused

This item is considered as poor item. It has approriate discriminating index but it also shows poorly difficulty index and one ineffective distractor.

Discussion item number 29

29. We can find the notice near a ...

a. Library   c. Pond

b. School   d. Hill

No Item	Opt	UPPER	LOWER
29	A	1	5
	B	0	3
	C*	23	11
	D	3	8
Total		27	27

1) The difficulty level (P) is 0.63. It means that the item is categorize into satisfactory difficulty level since only 34 students who get the right answer.

2) The discriminating index (D) is 0.44. It discriminates upper and lower group students in satisfactory level.

3) All the distractors are fungtioning well since these attract more students in lower group. The distractor analysis as follow:

a) Distractor A = acceptable

b) Distractor B = acceptable

c) Distractor D = acceptable

From the difficulty level, discriminating indexand distraction evaluation, this item can be categorized into good item.

I have a pet. It is a cat. Its name is Pussy. .... (30) has brown fur and a long tail. Its eyes .... (31) bright. It's so funny. It ....(32) rice and fish everyday. I love Pussy very much

Discussion item number 30

30. a. I   c. He

    b. It   d. She

No Item	Opt	UPPER	LOWER

30	A	0	6
	B*	23	8
	C	1	3
	D	3	10
Total		27	27

1) The difficulty level (P) is 0.57. It means that the item is categorize into satisfactory difficulty level since only 31 students who get the right answer.

2) The discriminating index (D) is 0.56. It discriminates upper and lower group students in satisfactory level.

3) All the distractors are fungtioning well since these attract more students in lower group. The distractor analysis as follow:

a) Distractor A = acceptable

b) Distractor C = acceptable

c) Distractor D = acceptable

From the difficulty level, discriminating indexand distraction evaluation, this item can be categorized into good item.

Discussion item number 31

31. a. Is   c. Are

b. am   d. Was

No Item	Opt	UPPER	LOWER
31	A	14	7
	B	0	3
	C*	6	5
	D	7	12
Total		27	27

1) The difficulty level (P) is 0.20. It means that the item is categorize into difficult level since only 11 students who get the right answer.

2) The discriminating index (D) is 0.04. It discriminates upper and lower group students in poor level.

3) Distractors B and D are fungtioning well since these attract more students in lower group. A is ineffective distractor because it attract more students in the upper group. The distractor analysis as follow:

a) Distractor A = revised

b) Distractor B = acceptable

c) Distractor D = acceptable

Although it has two effective distractor, it has poor difficulty level and discriminating index, so it is considered as poor item.

Discussion item number 32

32. a. Eats　　　　　　　　　　　c. Is eating

　　b. Eat　　　　　　　　　　　d. Ate

No Item	Opt	UPPER	LOWER
32	A*	8	3
	B	3	7
	C	10	10
	D	6	7
Total		27	27

1) The difficulty level (P) is 0.20. It means that the item is categorize into difficult level since only 11 students who get the right answer.

2) The discriminating index (D) is 0.19. It discriminates upper and lower group students in poor level.

3) Distractor B and D are functioning as intended because it attracts more students from lower group. Distractor   C is functioning as intended

because it attracts some students from lower and upper group in equal number. This is the analysis of the distractor:

a) Distractor B = acceptable

b) Distractor C = revised

c) Distractor D = acceptable

This item is considered as poor item. It has inappropriate difficulty level and it has poor discriminating index although most of the distractors are functioning well.

I ....(33) to Germany with my family ....(34) a school holiday a few years ago. It was getting dark, so we looked for a hotel. Then, we ....(35) upon two hotels next to each other with park. We just randomly chose and ....(36) there for the night.

Discussion item number 33

33. a. Go   c. Went

b. Am going   d. Was going

No Item	Opt	UPPER	LOWER
33	A	1	10
	B	3	10
	C*	20	6
	D	3	1
Total		27	27

1) The difficulty level (P) is 0.48. It means that the item is categorize into satisfactory difficulty level since only 26 students who get the right answer.

2) The discriminating index (D) is 0.52. It discriminates upper and lower group students in good level.

3) Distractor A and B are functioning as intended because it attracts more students from lower group. Distractor D is not functioning as intended because it attracts some students from upper group. This is the analysis of the distractor:

a) Distractor A = acceptable

b) Distractor B = acceptable

c) Distractor D = revised

This item is considered as good item. It has appropriate level of difficulty and discriminating index. It also has two effectiveness distractors.

Discussion item number 34

34. a. For   c. Since
    b. From   d. During

No Item	Opt	UPPER	LOWER
34	A	10	8
	B	9	15
	C	5	3
	D*	3	1
Total		27	27

1) The difficulty level (P) is 0.07. It means that the item is categorize into difficult level since only 4 students who get the right answer.

2) The discriminating index (D) is 0.07. It discriminates upper and lower group students in poor level.

3) Distractor A and C are not functioning as intended because it attracts more students from upper group. Distractor C is functioning as intended because it attracts some students from lower group. This is the analysis of the distractor:

a) Distractor A = revised

b) Distractor C = revised

c) Distractor D = acceptable

This item is considered as poor item. It has inappropriate difficulty level and it has poor discriminating index. Most of the distractors are not functioning well.

Discussion item number 35

35. a. Come   c. Comes

   b. Came   d. Coming

No Item	Opt	UPPER	LOWER
35	A	1	12
	B*	23	7
	C	1	3
	D	2	5
Total		27	27

1) The difficulty level (P) is 0.56. It means that the item is categorize into satisfactory difficulty level since only 30 students who get the right answer.

2) The discriminating index (D) is 0.59. It discriminates upper and lower group students in good level.

3) All of the distractor are functioning as intended because it attracts more students from lower group. This is the analysis of the distractor:

a) Distractor A = acceptable

b) Distractor C = acceptable

c) Distractor D = acceptable

This item considered as good item. It can fulfill three criteria of good test. It has appropriate difficulty level and discriminating index. It also has three distractors which function effectively.

Discussion item number 36

36. a. Stay   c. Are staying

   b. Stayed   d. Were staying

No Item	Opt	UPPER	LOWER
36	A	0	3
	B*	25	12
	C	2	9
	D	0	3
Total		27	27

1) The difficulty level (P) is 0.69. It means that the item is categorize into satisfactory difficulty level since only 37 students who get the right answer.

2) The discriminating index (D) is 0.48. It discriminates upper and lower group students in good level.

3) All of the distractor are functioning as intended because it attracts more students from lower group. This is the analysis of the distractor:

a) Distractor A = acceptable

b) Distractor C = acceptable

c) Distractor D = acceptable

This item can fulfill three criteria of good test. It has appropriate difficulty level and discriminating index. It also has three distractors which function effectively. This item considered as good item.

Discussion item number 37

37. Arrange these jumbled words into a good sentence

   <u>this</u> - <u>not</u> – <u>park</u> – <u>do</u> – <u>step</u> – <u>the</u> – <u>on</u>

    1   2   3   4   5   6   7

   a. 5-2-6-7-3-8-1-4   c. 7-4-5-2-6-8-1-4

b. 5-2-4-7-3-8-1-6   d. 1-4-5-2-6-8-7-3

No Item	Opt	UPPER	LOWER
37	A*	13	4
	B	2	5
	C	2	6
	D	10	12
Total		27	27

1) The difficulty level (P) is 0.31. It means that the item is categorize into satisfactory difficulty level since only 17 students who get the right answer.

2) The discriminating index (D) is 0.44. It discriminates upper and lower group students in good level.

3) All of the distractor are functioning as intended because it attracts more students from lower group. This is the analysis of the distractor:

a) Distractor B = acceptable

b) Distractor C = acceptable

c) Distractor D = acceptable

This item considered as good item. It can fulfill three criteria of good test. It has appropriate difficulty level and discriminating index. It also has three distractors which function effectively.

Discussion item number 38

38. Arrange these jumbled words into a good sentence

Easily - a box – yesterday – Fahmi – in the class – moved

  1   2   3   4   5   6

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

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

No Item	Opt	UPPER	LOWER
38	A	22	7
	B*	4	2
	C	1	12
	D	0	6
Total		27	27

1) The difficulty level (P) is 0.11. It means that the item is categorize into difficult level since only 6 students who get the right answer.

2) The discriminating index (D) is 0.07. It discriminates upper and lower group students in poor level.

3) Distractor A is not functioning as intended because it attracts more students from upper group. Distractor C and D are functioning as intended because it attracts more students from lower group. This is the analysis of the distractor:

a) Distractor A = revised

b) Distractor C = acceptable

c) Distractor D = acceptable

This item is considered as poor item. It has inappropriate difficulty level and it has poor discriminating index. Although most of the distractors are functioning well.

Discussion item number 39

39. Arrange these sentences to make a good paragraph!

1. They left at 7 a.m and arrived at the lake at 9 a.m.

2. After having lunch, they went home

3. There were 40 students all together

4. Their teacher, Miss Heny accompanied them to go

5. Last Sunday, the studentsof Santi's class had a trip to Singkarak Lake

6. They looked very happy playing in the lake

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

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

No Item	Opt	UPPER	LOWER
39	A	4	8
	B*	20	8
	C	3	8
	D	0	3
Total		27	27

1) The difficulty level (P) is 0.52. It means that the item is categorize into satisfactory difficulty level since only 28 students who get the right answer.

2) The discriminating index (D) is 0.44. It discriminates upper and lower group students in good level.

3) All of the distractor are functioning as intended because it attracts more students from lower group. This is the analysis of the distractor:

a) Distractor A = acceptable

b) Distractor C = acceptable

c) Distractor D = acceptable

This item considered as good item. It can fulfill three criteria of good test. It has appropriate difficulty level and discriminating index. It also has three distractors which function effectively.

Discussion item number 40

40. Arrange these sentences to make a good paragraph!

1. Finally, the fox look it and ran away

2. Long time ago, there was a crow

3. Suddentlythe fox came up and asked for her cheese

4. Then, the fox said that she had beautiful voice

5. One day, she ate a big piece of cheese

6. At the moment she opened her beak, the cheese fell over

7. However, shedid not give it to the fox

8. After that the crow sang a song

a. 2-5-3-7-4-8-6-1   c. 2-4-1-3-7-8-5-6

b. 2-6-7-3-8-5-4-1   d. 2-3-7-8-1-4-5-6

No Item	Opt	UPPER	LOWER
40	A*	23	13
	B	1	6
	C	2	5
	D	1	3
Total		27	27

1) The difficulty level (P) is 0.67. It means that the item is categorize into satisfactory difficulty level since only 4 students who get the right answer.

2) The discriminating index (D) is 0.37. It discriminates upper and lower group students in satisfactory level.

3) All of the distractor are functioning as intended because it attracts more students from lower group. This is the analysis of the distractor:

a) Distractor A = acceptable

b) Distractor C = acceptable

c) Distractor D = acceptable

   This item considered as good item. It can fulfill three criteria of good test. It has appropriate difficulty level and discriminating index. It also has three distractors which function effectively.

Discussion item number 41

41. Complete the text with the correct answer!

Faisal and Ali are twins. They look the same, but they have some differences. Faisal is 39 kg, but Ali is 47 kg. So, Ali is (a).....than Faisal. Ali is 160 cm tall, but Faisal is 163 cm tall. So, Faisal is (b)....than Ali. Faisal always passes the math tests. Ali sometimes fails in the math tests. So, Faissal is (c).... at math than Ali. Both Faisal and Ali like to collects comics. Faisal has 25 comics and Ali has 30 comics. So Ali has (d).... comics than Faisal.

1) The difficulty level (P) is 0.61. It means that the item is categorize into satisfactory difficulty level
2) The discriminating index (D) is 0.17. It discriminates upper and lower group students in poor level.

This item considered as good item. It has appropriate difficulty level but poor discriminating index. This item need to be revised, before used again.

Discussion item number 42

42. Change the word in the braket into the correct order!

Last Sunday, my family and I (go) to the Parangtritis beach. Arrived there, we (swim) in the beach. Then, we (play) sand. We came home at night. We (be) happy.

1) The difficulty level (P) is 0.62. It means that the item is categorize into satisfactory difficulty level
2) The discriminating index (D) is 0.15. It discriminates upper and lower group students in poor level.

This item considered as good item. It has appropriate difficulty level but poor discriminating index. This item need to be revised, before used again.

Discussion item number 43

43. Write a short message to your friend that you will come to his/her house to borrow his/her Englis book this afternoon.

   1) The difficulty level (P) is 0.62. It means that the item is categorize into satisfactory difficulty level.
   2) The discriminating index (D) is 0.33. It discriminates upper and lower group students in satisfactory level.

This item considered as good item. It has appropriate difficulty level and discriminating index.

Discussion item number 44

44. Suppose you are the scout leader, write an announcement to all the scout member that there is no scout activity during the test. it will start again after the test.
   1) The difficulty level (P) is 0.56. It means that the item is categorize into satisfactory difficulty level.
   2) The discriminating index (D) is 0.33. It discriminates upper and lower group students in satisfactory level.

      This item considered as good item. It has appropriate difficulty level and discriminating index.

Discussion item number 45

45. Rearrange the sentences bellow into good order!
   a. His short bill could not reach the water
   b. He looked everywhere to find some water, but there was none
   c. Then,he found a tall jar with very little water in it
   d. At last, he drank the water, as much as he liked
   e. Soon, he filled the jar with stones so the water came up higher
   f. One day, a crowwas very tired and thirsty

1) The difficulty level (P) is 0.56. It means that the item is categorize into satisfactory difficulty level since only 4 students who get the right answer.

2) The discriminating index (D) is 0.22. It discriminates upper and lower group students in satisfactory level.

This item considered as good item. It has appropriate difficulty level and discriminating index.

CHAPTER V

CONCLUSION AND SUGGESTION

A. Conclusion

After analyzing the data, the conclusion deduced as follow:

1. The level of difficulty of English final test for eighth grade students of MTs N Klaten was 25% multiple choice items included easy items, 65% multiple choice items and 100% essay items included satisfactory items and 10% difficult items. It was dominated by satisfactory items. So that the test items were good enough for the students. It lead to have good of level difficulty.

2. The discrimination index of English final test for eighth grade students of MTs N Klaten were 32,5% multiple choice items and 40% essay items included poor items, 27,5% multiple choice items and 60% essay items included satisfactory items and 40% multiple choice items included good items. Both satisfactory and poor discrimination index items dominated this test. It meant that the discrimination index for each of items was balance.

3. The percentage of distractors for each items in English final test for eighth grade students of MTs N Klaten were 22,50% categorized as omit, 71,67% effective distractor, and 5,83% less effective distractor.

4. While the precentage of the over all data were 35% multiple choice items and 60% essay items include into accepted item. 37,5% multiple choice items and 40% essay items included revised items. And 27,5% multiple choice item included refused items.

On the basis of the conclusion above, it could be drawn a general conclusion that the quality of English final test for eighth grade students of MTs N Klatenin academic year 2016/2017 wasgoodenoughin term of boththe level of difficulty, discriminating indexand the effectiveness of

thedistractors.But overall, there are many items that need to be revised because it only fulfill 2 criterias.Those aspects of test need to be improved.

B. Suggestion

Based on the research study which has done, delivered the suggestion as follows:

1. The teachers hopefully improve their ability in making standardized test or in analyzing test items by using item analysis.

2. MGMP Team of English subject hopefully have good coordination in arranging test items and analyzing test items for wide scope in order to gained qualified test items and appropriate with standard.

3. The other researcher can use this research as well to conduct related research. For other researcher who want to develop such a research, this study might be used to be helpful reference.

REFERENCES

Arifin, Zainal. 2012. *Evaluasi Pembelajaran* . Bandung : Remaja Rosdakarya.

Arikunto, Suharsimi. 2013. *Dasar-dasar Evaluasi Pendidikan* . Jakarta: Bumi Aksara .

BSNP. *Panduan Penulisan Butir Soal*. Direktorat Pembinaan SMP Ditjen Menejemen Pendidikan Dasar dan Menengah Kementerian Pendidikan Nasional.

Brown, H. Douglas. 2003. *Language Assesment and Principle and Classroom Practice.* San Francisco: Addition Wesley Longman, Inc.

Cresswell, Jhon. 2012. *Educational Research Planning, Conducting and Evaluationg Qualitative and Quantitative Research.* USA: SAGE .

Fulcher, G. 2010. *Practical Language Testing.* United Kingdom: Hodder Education.

Gronlund, N. E. 1982. *Constructing Achievement Tests.* United State of America: Virginia Livsey.

Hanafi, Ahmad. Suhardi. 2014. *The Characteristics of the English Test Items in NE for VHS's in Tabalong Regency South Kalimantan in the academic year 2010/2011.* Jurnal LingTera, 1(1).

Harmer, Jeremy. 2001. *The Practice of English Language Teaching* . China: Pearson Education.

Hughes, Arthur. 2003. *Testing for Language Teaching* . New York: Cambridge University.

Louis Cohen, L. M. 2007. *Research Methods in Education 6th Edition* . United State of America: Routledge.

Nurliyanto, Dian. 2015. *Test Item Analysis of the Final Examination on Economics Subject in Grade XII IPS SMA Negeri Banyumas Academic year 2014/2015.* Thesis tidak diterbitkan. Yogyakarta: Fakultas Ekonomi Universitas Negeri Yogyakarta

Jordan, Elizabeth. 2006. *Educational Psychology A Problem-Based Approach* . Ubited State of America: Pearson.

Shih, Yu-mien. 2010. An Item Analysis of an English Achievement Test Taken by EFL College Students in Taiwan. *Taiwan Journal 6*.

Sangadji, Etta Mamang & Sopiah. 2010. *Metodologi Penelitian – Pendekatan Praktis dalam Penelitian*. Yogyakarta: Andi Offset.

Saputri, Ratih Widya. 2016. The Analysis of the English Mid-Term Test Made by the Teacher for Eighth Students of SMP Negeri 1 Wuryantoroat the Even Semesterof 2014/2015 Academic Year. State Islamic Institute of Surakarta

Sudijono, Anas. 2012. *Pengantar Evaluasi Pendidikan* . Jakarta: Raja Grafindo Persada.

Sukamolson, Suphat. *Fundamentals of Quantitative Research*. DOAJ.