AN ANALYSIS OF MULTIPLE CHOICE ITEMS OF ENGLISH FINAL SEMESTER TEST MADE BY ENGLISH TEACHER (Study at the Tenth Grade MA AL-Muthmainnah in Academic Year 2020/2021)

By ERI ANGGRIANI

WORD COUNT 7533 TIME SUBMITTED 28-JUL-2021 09:54AM

PAPER ID 74927486

AN ANALYSIS OF MULTIPLE CHOICE ITEMS OF ENGLISH FINAL SEMESTER TEST MADE BY ENGLISH TEACHER
(Study at the Tenth Grade MA AL-Muthmainnah in Academic Year 2020/2021)

A Thesis

Presented as A Partial Fulfillment of the Requirement for the Bachelor Degree in English Language teachingFaculty of Teacher training and Education Muhammadiyah University of Mataram

BY

ERI ANGGRIANI
NIM. 117120001

ENGLISH EDUCATION PROGRAM

FACULTY OF TEACHER TRAINING AND EDUCATION

MUHAMMADIYAH UNIVERSITY OF MATARAM

2021

CHAPTER I

INTRODUCTION

1.1 Background of the Study

Assessment is essentially critical in the teaching and learning process. The assessment also plays an important role. The assessment evaluates the students' progress throughout the learning process by gathering the information needed variously. It also provides information to the teacher on the learning outcomes that the students have achieved.

The infield of the teaching and learning process, especially in English, cares to know if we are doing the right thing and whether our activities will lead to the result we want to achieve. Thus, the evaluation is needed to measure student's achievements in the teaching dan learning process. The assessment itself is a process of assessing to describe student achievement according to established criteria. In this evaluation, the teacher will see the level of student achievement that who has achieved. Providing daily tests, midterms, and final exams is one of the educational evaluation programs. Given a test is a part of the evaluation. A test is usually defined as a collection of items whose answers can be true-false.

In MA AL-Muthmainnah, the assessment most often used by the teacher is multiple-choice. Thus, the measuring instrument used in the assessment must provide an overview of students' learning abilities. Therefore, it is

necessary to analyze the quality of the multiple-choice items. This question analysis aims to determine the level of difficulty of the items and thediscriminationpower of multiple-choice items. In the multiple-choice items, especially the English final semester, most of the teachers arranged the questions without paying attention to the quality of the multiple-choice items, one of which was the difficulty level and discriminating power of the items themselves.

At this moment, MA AL-Muthmainnah has never previously conducted research related to multiple-choice items of the English final semester. Based on the problem found at tenth grade MA AL-Muthmainnah, students' different levels of knowledge can be seen from their performance in the test.Thus, the researcher believes that analyzing the item is very important for creating a good test and improving test items to which extent the quality of a test can be determined. Not only for improving the quality of test items but also the analysis items are intended to determine to what extent the test clearly shows the difference among students regarding their level of knowledge of content and material of lessons being tested.

Seeing this fact, the researcher chooses MA AL-Muthmainnah as the place to conduct this research. MA AL-Muthmainnah is one of the private schools in the Soromandi sub-district. According to the researcher's information, the types of items used to assess students in the test were multiple-choice. With these facts, and the researcher decided to research at the school.

There are some researchers, conduct research about multiple-choice. The first is research conducted by AyuAmaliyah (2018) with the title "An analysis of multiple-choice items made by the teacher based on Bloom's revised taxonomy theory at tenth grade of SMA NegeriSidoarjo. Second, research was conducted by Mina, Regina, BambangWijaya (2013), and the title is "An analysis on the English multiple-choice item test for primary students". The previousresearch conducted by Hartati, N., & Yogi, H. P. S. (2019) the title is "item analysis for a Better Quality Test". So, the existence of these previous studies, the research was conducted related to multiple-choice to determine the test's quality in assessing and measuring students' achievement.

The phenomena above encourage the writer to do the descriptive analysis of multiple-choice items used for the final semester at tenth-grade students MA AL-Muthmainnah 2020/2021 academic year based on difficulty level and discriminating power of multiple-choice items. The researcher wants to research the title" An Analysis of Multiple choice items of Englishfinalsemesterat the tenth-grade MA AL-Muthmainnah in the academic year 2020/2021.

1.2 Research Question

Based on the background of the study above, the question is formulated as follow:

1. How is the difficulty level of the multiple-choice items of the final semester test at the tenth grade MA AL-Muthmainnah in the academic year 2020/2021?

2. How is the discrimination power of the multiple-choice items of the final semester test at the tenth grade MA AL-Muthmainnah in the academic year 2020/2021?

1.3 Research objectives

As seen from the research questions above, the research objectives are:

1. To find out the difficulty level of multiple-choice items of the final semester test at the tenth grade MA AL-Muthmainnah in the academic year 2020/2021.
2. To examine the discrimination power of multiple-choice items of final semester test at the tenth grade MA AL-Muthmainnah in the academic year 2020/2021.

1.4 Scope and Limitation of Research

This research has beenconducted in MA AL-Muthmainnah at the tenth-grade students. The study focused on multiple-choice items in the tenth grade in the final semester test. The subject of the study was studentsin the tenth grade in the MA AL-Muthmainnah, while the object of the study is the multiple-choice question sheet and students answer sheets from the tenth grade in the MA AL-Muthmainnah.

1.5 Definition of Key terms

1. Multiple-choice

   According to Heaton (1988: 30),A multiple-choice is a form of assessment where respondents were asked to choose the best answer.

2. Final Semester

   Assessment in education system national test consisting of daily tests, general test, classroom promotion tests, school final exams, and national exams (Depdiknas, 2003)

3. Teachers

   According to Suparlan (2008: 12), teachers can be interpreted as people whose duties are related to educating the nation's life in all aspects, includingspiritual and emotional, intellectual, physical, and other aspects.

CHAPTER II

REVIEW OF RELATED LITERATURE

2.1 Previous Research Findings

The researcher has some relevant previous researches that support this research. The final project: inspires the research

The first research is conducted by AyuAmaliyah (2018) by the title "An analysis of multiple-choice items made by the teacher based on Bloom's revised taxonomy theory at tenth grade of SMA Negeri 1 Sidoarjo. This research was conducted; this study aims to know levels of learning in multiple-choice items made by the teacher based on the new version of Bloom's cognitive level at tenth grade of SMA Negeri 1 Sidoarjo. This study uses a descriptive qualitative approach. To collectedthe data, this research used documentation. The researcher analyzed 35 multiple-choice items using the new version of Bloom's cognitive level.

The second research was conducted by Mina, Regina, BambangWijaya (2013) with the title "An analysis on the English multiple-choice item test for primary students". This research concentrates on specific questions regarding the item validity, the test scores reliability, and item analysis to provide information that will improve test items construction. A descriptive method is applied to describe and examine the data.

The last research was conducted by Sartati, N., & Yogi, H. P. S. (2019) entitle "Items Analysis for a Better Quality Test". This study is the small-scale study of item analysis of a teacher's own-made summative test. It examines

the quality of multiple-choice items regarding the difficulty level, the discriminating power, and the effectiveness of distractors. The study employed a qualitative approach which also used a simple quantitative analysis to analyze the quality of the test items through the document analysis of the teacher's English summative test and the students' answer sheets.

In addition, the similarities and differences between my research and the three theses above are: the first thesis by AyuAmaliah (2018) is the same as analyzing multiple-choice items, and the difference is that this thesis uses a revised taxonomy theory to analyze the multiple-choice items, whereas I don't use any theory, I only focus on analyzing the level of difficulty and discriminating power on the multiple-choice question. In the second thesis by Mina, Regina, BambangWijaya (2013), the similarities with this thesis is the same as analyzing multiple-choice items. Still, this thesis only focuses on difficulty level and discriminating power. At the same time, the only difference in this thesis is that this thesis analyzes selected questions in primary students. In contrast, the writer analyzes the multiple-choice items at the senior high school level, for the last journal by Hartati, N., & Yogi, H. P. S. (2019), the similarities this research, that the writer concerning on difficulty level and discriminating power of multiple-choice items. In contrast, the difference in this journal is the item analysis of the English summative test made by the teacher.

2.2 Theoretical Review

1. Test

a. Definition of Test

Tests are questions that are given to students to get answers from students in oral form (oral test), in written form (written test), or in the form of action (test action) (Sudjana, 2013: 35). According to Uno and Koni (2012: 111) test, a set of stimuli is given to someone to get the answer to be the basis for scoring. The score is based on a score representative of test follower behavior, which indicates the extent to which the test person has moderate characteristics. Furthermore, Koyan (2011: 7) states that the test is the instrument or systematic tool which consists of set questions or tasks to measure a particular behavior in learners with the help of a numeric scale or specific categories.

From the description of the experts above, it can be concluded that the test is a systematic way or tool to measure student ability, in which there is a set of questions to get answers from students either in oral, written, or deed form and used as a basis for assessment and scoring of numbers.

Heaton (1988: 5) stated that tets might be constructed primarily as devices to reinforce learning and motivate the student or assess the student's performance in the language. In the former case, the test is geared to the teaching that has taken place, whereas in the latter case, the teaching is often geared main 6 he test. Then, Brown (2004: 3) added

that a test is a method of measuring a person's ability or knowledge in a given domain. It consists of some systematic procedures for gathering data about students' achievement and can be instruments, techniques, or procedures to have the students respond through performance or tasks in the form of a set question.

Hughes (1993: 7) stated that testing has several purposes, as follows:

1. To measure language proficiency.
2. To discover how successful students have been in achieving the objectives of a course of study.
3. To diagnose students' strengths and weaknesses, to identify what they know and what they do not know;
4. To assist students' placement by identifying the stage or part of a teaching program most appropriate to their ability.

2. Types of Test

The re are some types of tests, according to Hughes (1993:9). There are:

1. Proficiency Test

Proficiency tests are tests designed to measure people's ability in a language, regardless of any training they may have had in that language. Therefore, the content of a proficiency test is not based on the content or objectives of language courses that people taking the test may have

followed. Rather, it is based on a specification of what candidates have to be able to do in the language to be considered proficient.

2. Achievement Test

As seen from the name, the purpose of this achievement test is to determine the success of students, both individually and in groups, or the courses themselves have been in achieving objectives. Brown (2004:47) stated that an achievement test is related directly to classroom lessons, units, or even an entire curriculum.

It can be concluded that the learning outcome test is used to assess the level of student success in learning that has been determined following the objectives of the learning. Teachers also use learning outcomes tests to motivate students to study hard. For example, by giving quizzes every week or at the end of the semester, the effect is often an increase in study time near the test of time.

3. Aptitude Test

Brown (2004:43) stated that an aptitude test is designed to measure capacity or general ability to learn a foreign language and ultimate success in that undertaking. Aptitude tests are ostensibly designed to apply to the classroom learning of any language. Ormrod (2008:347) also stated that an aptitude test is to identify students who are most likely to succeed in certain subjects. The test may also be used to counsel students about educational plans and career choices in the future.

This aptitude test measures students' ability to learn so they know specific talents of students, for example, in language learning, whether students have special talents

a. Diagnostic Test

Heaton (1988:173) stated that diagnostic test is widely used; few tests are constructed solely as diagnostic tests. Note that diagnostic testing is frequently carried out of groups of students rather than for individuals.

b. Placement Test

Brown (2004:54) stated that certain proficiency tests could act in the role of placement tests, the purpose of which is to place a student into a particular level or section of a language curriculum or school. A placement test usually, but not 11 always, includes a sampling of the material to be covered in the various course in a curriculum, a students performance on the test should indicate the point at which the students will find material neither too easy nor too difficult but appropriately challenging

3. Item Analysis

A good test must also be good at its item analysis; there are several simple statistical ways to check individual items. Brown (2004:58-60) stated that "they are: difficulty level, discriminating power, and distractor effectiveness.

a. The Difficulty level of the item

The difficulty level is one of a kind of item analysis. The level of difficulty was concerned with how difficult or easy the item for the students. Shohamy (1985:79) states that difficulty level relates to how easy or difficult the item is from the students who took the test. It is essential since test items that are too easy can tell us nothing about the test population's differences. If the item is too easy, most or all of the students obtained the correct answer.

In contrast, if the item is difficult, most or all of the students get it wrong. The quality of options is a distribution of tests in decided alternatives on a multiple-choice test. It is obtained by calculating the number of oftest who choose the alternatives A, B, C, or D or those who do not choose any alternatives. In this way, the teachers would be able to identify whether distractors function well or badly.

Arikunto (2013: 222) stated that a good test is not too easy or too difficult. A good test item must have a difficulty level, including easy, medium, and difficult levels. Moderate level items belong to the category of effective and good tests. Items that are too easy or difficult could weaken the test's quality, and the validity of the information about data Student achievements will not be acquired.

The difficulty index (P) of the itemscan be determined by calculating the proportion of test-takers that answer the item correctly. The following is the formula for calculating the item difficulty index:

$$P=NP/N$$

In which:

P = Indeks of difficulty level

NP = Number of test-takers answering correctly

N = number of test-takers responding to that item.

(Bachman, 1990:125)

In this case, the difficulty level must be classified on the rank scale of difficulty level as follows:

Table 2.1 The Classification of item Difficulty Level

Interpretation	P
Difficult	0,00 – 0,30
Moderate	0,31 – 0,70
Easy	0,71 – 1,00

(SuharsimiArikunto, 2013: 210)

b. Discrimination Power

Brown (2004:59) stated that discriminating power is the extent to which an item differentiates between high and low test-takers. Sudijono (2012: 387) stated that discriminating is calculated based on classification into two groups, higher and lower groups.

Item discrimination index (D) can be obtained by dividing into two groups according to their scores on the test as a whole: a higher group and a lower group. The following formula is employed to determine the item discrimination index:

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B}$$

Where:

D: Item Discrimination (Discrimination Power)

BA: number of top test takers that have the correct answer

BB: number of bottom test takers that have the correct answer

JA: total participant of top test-takers

JB: total participant of bottom test-takers

(Brown, 2004: 59)

After the item discrimination index is found, the discriminating power of an item can be determined. The determine the discriminating power, and the following classification can be used to indicate whether the discriminating power of an item is excellent, good, satisfactory, poor or worst.

Table 2.2 The Classification of Item Discriminating Power:

Discrimination Index	Quality
0,00 – 0,20	Poor
0,21 – 0,40	Satisfactory
0,41 – 0,70	Good
0,71 – 1.00	Excellent

(SuharsimiArikunto. Op. Cit., p. 232)

c. Distractor Analysis

Furthermore, the last item analysis activity is effectiveness bully. This is a procedure that deals specifically with multiple-choice items. This switch's function is to divert students from the correct answer if they do not know correctly. According to Sudijono (2012: 411), determining whether a distractor 17 can work effectively if a switch has been selected by at least 5% of all test takers. The effectiveness of distractor analysis provides information about how successfully a

distraction has distracted students who have not learned well from the correct answers.

4. Multiple Choice items

According to Heaton (1988: 30),A multiple-choice is a form of assessment where respondents were asked to choose the best answer. This item is most frequently used in educational testing. According toPopham (2011: 148) stated that the multiple-choice items could be used to measure a student's possession of knowledge or the student's ability to engage in higher thinking levels.

AccordingtoSJ Burton (1991) stated that The advantages of multiple-choice items are:

1. Multiple-choice items are objectively scored; they are not affected by scorer inconsistencies, as are essay questions.
2. Therefore, a student is able to answer many multiple-choice items in the time it would take to answer a single essay question. This feature enables the teacher to use multiple-choice items to test a broader sample of course content in a given amount of testing time.
3. Multiple-choice items are amenable to rapid scoring, which is often done by scoring machines.

Heaton (1988) stated that The general principles should be observed when multiple-choice items are constructed:

1. Each multiple-choice item should have only one answer.

2. Only one feature should be tested: it is usually less confusing for the tester, and it helps to reinforce a particular teaching point. Few would wish to test both grammar and vocabulary simultaneously, but sometimes word order and sequence of tenses are tested simultaneously.
3. Each option should be grammatically correct when placed in the stem, except, of course, in the case of specific grammar test items.
4. All multiple-choice items should be at a level appropriate to the proficiency level of the testers.
5. Multiple-choice items should be as brief and as clear as possible (though it is desirable to provide short context for grammar items).
6. In many tests, items are arranged in rough order of increasing difficulty. It is generally considered essential to have one or two simple items to "led in" the testers, especially if they are not too familiar with the kind of test being administered

CHAPTER III

RESEARCH METHODOLOGY

In this chapter, the researcher talked about research methods. Research methods are indispensable in conductedresearch; with this helping, the researcher found their study results. Research methods related to how we do research are constructive for researchers in solving problems. Also, the subject matter will be straightforward and easy to understand if we apply adequate methods. This chapter will describe the steps, namely, method of study, research location, data and source data research instrument, data collection technique and data analysis technique.

3.1 Method of the Study

This research was a case study.In Creswell, 2010: 20). Stake (in Denzim& Lincoln, 1994: 236-238) details the characteristics of the study cases are as follows: 1). Case studies are a form of research (inquiry) or the study of a specific problem (particularity). 2). It can be done either with a qualitative approach or quantitative, but more emphasis is placed on a qualitative approach. 3). Case study objectives can be in the form of individuals or groups, even the wider community.

This study is categorized as quantitative research using descriptive analysis. The writer described the difficulty level and discrimination power by analyzing the multiple-choice items of the English final semester of the tenth-grade student's MA AL-Muthmainnah in the academic year 2020/2021.

3.2 Data and Source of Data

The researcher conducted this Research at MA Al-Muthmainnah. The data and the source of this research was the document analysis. The data and source of data were documented. The document came from the multiple-choice items sheet and students' answer sheets. Then the research instrument wasdocument analysis. There were 25 multiple-choice items and 16 students took the final semester test at the tenth grade in MA AL-Muthmainnah.

3.3 The Technique of Data Collection

This research data was collected through the documentation study.Datais obtained from 16 of the students' answers to items in the English final semester test. A total of 25 items in the form of multiple-choice were used in the item analysis process. For the scoring technique, every correct answer is given one point and zero for each wrong answer.

There are several steps in data collection techniques, namely:

1. The writer came to the school and introduced the writer to the school residents.
2. After that,the writer came to the classroom. Then introduced to the English teacher and students.
3. The writer took the question sheet sample/document from the English teacher.
4. The last step was analyzing the multiple-choice test by ficulty level and discrimination power of multiple-choice items.

3.4.Data Analysis Technique

For data analysis, descriptive statistics were employed in this research. In a descriptive quantitative approach, descriptive statistics are defined as numerical, graphical, and tabular techniques for organizing, analyzing, and presenting data (George Argyrous, 2011: 20). This research described the difficulty level and discriminating power for each multiple-choice test item based on the index.

In addition, the steps for analyzing the data were as follow:

1. Taking the English multiple-choice paper and students' answer sheets.
2. Arranged or ranked the score from the highest to lowest
3. Divided the students into two groups: upper and lower groups by arranging the students' scores from the highest to the lowest.
4. Analyzing the test based on the difficulty and discriminating power
5. Interpreting the level of difficulty and classification of discriminating power based on the numerical data.
6. The writer was describing the difficulty and item discriminating power of the English multiple-choice items test.

To determine the difficulty level and discriminating index of items, the data has beencalculated by using the following formula:

The formula of item difficulty index:

P=NP/N

P = Indeks of difficulty level

NP = Number of test-takers answering correctly

N = number of test-takers responding to that item.

(Bachman, 1990:125)

The formula of item discrimination index:

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B}$$

Where:

D: Item Discrimination (Discrimination Power)

BA: number of top test takers that have the correct answer

BB: number of bottom test takers that have the correct answer

JA: total participant of top test-takers

JB: total participant of bottom test-takers

(Brown, 2004: 59)

For interpreting the level of item difficulty and discriminating power.The difficulty level could be found out by the classification of difficulty level indeks as follows:

Table. 3.1 The Difficulty Level Classification

Interpretation	P
Difficult	0,00 – 0, 30
Moderate	0,31 – 0,70
Easy	0,71 – 1,00

(SuharsimiArikunto, 2013: 210)

Table 3.2 The Classification of Item Discriminating Power:

Discrimination Index	Quality
0,00 – 0,20	Poor
0,21 – 0,40	Satisfactory
0,41 – 0,70	Good
0,71 – 1.00	Excellent

(SuharsimiArikunto, 2016p. 232)

CHAPTER IV

FINDING AND DISCUSSION

In this part, the researcher discusses the finding and discussion the research results. The English teacher's analysis of multiple-choice items of the English final semester test at the tenth grade MA AL-Muthmainnah in the academic year 2020/2021 is calculated statistically.

4.1 Finding

In this study, the researcher used the English final semester test made by the English teacher at the tenth-grade students of MA AL-Muthmainnah in the academic year 2020/2021 and student answer sheets as the data sources. As for the number of test items, namely all multiple-choice items. The data for this study were collected from 16 student answer sheets. A total of 20 items in the form of multiple-choice were analyzed in the item analysis process.

The data were analyzed statistically and manually calculated to determine the item difficulty index and the discrimination index. Then, index the item difficulty and discrimination of each test item were found in the numerical data (table and chart form).

4.1.1 Item difficulty Level

In analyzing the difficulty level, first, the researcher made a list of students' answers into the table. Then researcher calculated the level of difficulty with the following formula:

$$P=NP/N$$

In which: P = Indeks of difficulty level

NP = Number of test-takers answering correctly

N = number of test-takers responding to that item.

(Bachman, 1990:125)

After the difficulty level is calculated, the following, Then the researcher classified the results of the difficulty level the item is considered difficult, moderate, or easy based on this range scale:

Table 4.1 The Classification of item Difficulty Level

Interpretation	P
Difficult	0,00 – 0,30
Moderate	0,31 – 0,70
Easy	0,71 – 1,00

(SuharsimiArikunto, 2013: 210)

Analysis of the difficulty level of the multiple-choice items made by the English teacher at the tenth grade MA AL-Muthmainnah, There are 25 item numbers for the English final semester test. The data shows that there were 17 (68%) moderate items, 6 (24%) difficult items, and there were 2 (8%) easy items. The following the description of the statistically calculated data:

		butir soal																								
No	Nama	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	total
1	Sukmayati	0	1	0	0	1	1	1	0	0	1	1	1	1	0	0	1	0	0	0	1	0	1	0	1	0	12
2	Sisa Anggraini	1	1	0	0	1	0	0	1	0	0	1	0	1	0	1	0	0	0	1	0	0	1	0	0	1	10
3	Nurmajidah	0	0	1	0	0	1	1	1	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	1	1	9
4	Asrul Rahman	1	1	0	0	1	0	1	1	0	0	0	1	1	0	0	1	1	0	1	0	1	1	0	0	0	12
5	Ainun Dwi Rahum	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1	1	0	1	7
6	Nurul Fatiah	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0	5
7	Muzakir	1	0	1	1	0	1	0	0	0	1	0	1	1	0	1	1	1	0	1	0	1	0	0	0	1	14
8	Haikal	0	1	0	0	0	1	1	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	1	7
9	Siti Kirani Fadilla	0	1	1	1	1	0	1	0	0	0	1	0	0	1	1	0	1	1	0	1	0	1	0	0	1	13
10	Nurlaila	0	1	0	0	1	0	1	1	0	0	1	1	0	0	0	1	0	0	1	0	0	1	0	1	0	10
11	Emilia	1	0	0	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	6
12	Raisya Madani	1	1	0	1	1	0	1	0	1	1	0	0	1	0	0	0	1	0	1	0	0	1	0	0	1	12
13	Putri Ainun Hikmah	1	0	1	0	1	1	1	0	1	1	1	0	1	1	1	1	1	1	0	1	0	1	1	0	1	18
14	Lydia	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	1	8
15	Lira Fima	1	1	1	1	1	0	1	0	1	0	1	0	1	0	0	0	1	1	0	1	1	0	0	0	1	14
16	Tita Anjalita	0	1	1	0	1	0	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	0	1	13
	jumlah	9	11	6	5	11	5	10	7	3	6	6	4	10	4	7	7	10	5	6	6	2	12	3	3	12	
	level difficulty	0.56	0.69	0.38	0.31	0.69	0.31	0.63	0.44	0.19	0.38	0.38	0.25	0.63	0.25	0.44	0.44	0.63	0.31	0.38	0.38	0.13	0.75	0.19	0.19	0.75	

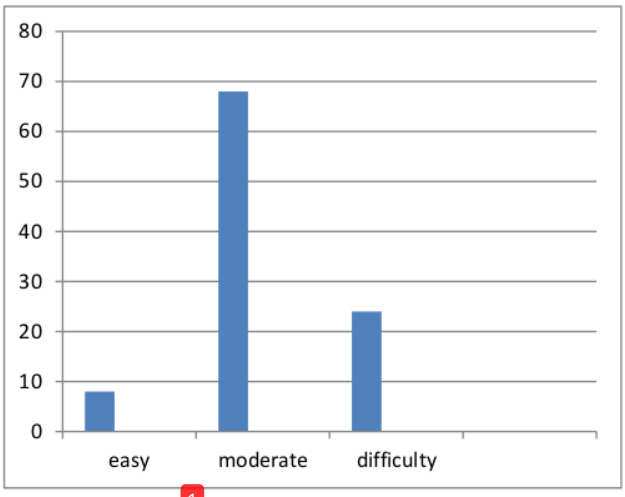
Each multiple-choice items is interpreted into a classification table to read the information easier about the difficulty level, as follows:

Table 4.2.The difficulty level analysis.

Item Number	P	Classification
1	0,56	Moderate
2	0,69	Moderate
3	0,38	Moderate
4	0,31	Moderate
5	0,69	Moderate
6	0,31	Moderate
7	0,63	Moderate
8	0,44	Moderate
9	0,19	Difficult
10	0,38	Moderate
11	0,38	Moderate
12	0,25	Difficulty
13	0,63	Moderate
14	0,25	Difficult
15	0,44	Moderate
16	0,44	Moderate
17	0,63	Moderate
18	0,31	Moderate
19	0,38	Moderate
20	0,38	Moderate
21	0,13	Difficult
22	0,75	Easy
23	0, 19	Difficult
24	0,19	Difficult
25	0,75	Easy

As seen from the table above, there were three columns in the table: the first column showed the number of items. The second column showed the difficult level analysis of items, and the third column showed the Classification of the difficulty level.

From the table above, it could be seen that the moderate item categorized there were 17 (68%) items that are number 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, and 20. The difficult items were 6 (24%) items that are number 9, 12, 14, 21, 23, 24, and two (8%) easy items that are 22 and 25.



For more details, the researcher describes each item as follows:

1. Item number 1 is moderate because of 16 students, nine students can answer correctly, and the difficulty level of this item is 0,56, included in the moderate item category.
2. Item number 2 is moderate because of 16 students who can answer correctly, there were 11 students, and the difficulty level of this item is 0,69, included in the moderate item category.
3. Item number 3 is moderate because there were six students from 16 students who can answer correctly, and the difficulty level of this item is 0,38 included in the moderate item category.

4. Item number 4 is moderate because there were five students from 16 students who can answer correctly, and the difficulty level of this item is 0,31 included in the moderate item category.
5. Item number 5 is moderate because there were 11 students from 16 students who can answer correctly, and the difficulty level of this item is 0,69, included in the moderate item category.
6. Item number 6 is moderate because there were five students from 16 students who can answer correctly, and the difficulty level of this item is 0,31 included in the moderate item category.
7. Item number 7 is moderate because there were ten students from 16 students who can answer correctly, and the difficulty level of this item is 0,63 included in the moderate item category.
8. Item number 8 is moderate because there were seven students from 16 students who can answer correctly, and the difficulty level of this item is 0,44 included in the moderate item category.
9. Item number 9 is difficult because there were 3sudents from 16 students who can answer correctly, and the difficulty level of this item is 0,19 included in the difficulty item category.
10. Item number 10 is moderate because six students from 16 can answer correctly, and the difficulty level of this item is 0,38, included in the moderate item category.

11. Item number 11 is moderate because six students from 16 can answer correctly, and the difficulty level of this item is 0,38, included in the moderate item category.
12. Item number 12 is difficult because of 16 students, four students can answer correctly, and the difficulty level of this item is 0,25 included in the difficulty item category.
13. Item number 13 is moderate because there were ten students from 16 students who can answer correctly, and the difficulty level of this item is 0,63 included in the moderate item category.
14. Item number 14 is difficult because there four students from 16 students who can answer correctly, and the difficulty level of this item is 0,25 included in the difficulty item category.
15. Item number 15 is moderate because there were seven students from 16 students who can answer correctly, and the difficulty level of this item is 0,44 included in the moderate item category.
16. Item number 16 is moderate because there were seven students from 16 students who can answer correctly, and the difficulty level of this item is 0,44 included in the moderate item category.
17. Item number 17 is moderate because there were ten students from 16 students who can answer correctly, and the difficulty level of this item is 0,63 included in the moderate item category.

18. Item number 18 is moderate because there were five students from 16 students who can answer correctly, and the difficulty level of this item is 0,31 included in the moderate item category.
19. Item number 19 is moderate because there were six students from 16 students who can answer correctly, and the difficulty level of this item is 0,38 included in the moderate item category.
20. Item number 20 is moderate because there were six students from 16 students who can answer correctly, and the difficulty level of this item is 0,38 included in the moderate item category.
21. Item number 21 is difficult because there were two students from 16 students who can answer correctly, and the difficulty level of this item is 0,13 included in the difficult item category.
22. Item number 22 is easy because there were 12 students from 16 students who can answer correctly, and the difficulty level of this item is 0,75 included in the easy item category.
23. Item number 23 is difficult because there were three students from 16 students who can answer correctly, and the difficulty level of this item is 0,19 included in the difficult item category.
24. Item number 24 is difficult because there were three students from 16 students who can answer correctly and the difficulty level of this item is 0,19 included in the difficult item category.

25. Item number 22 is easy because there were 12 students from 16 students who can answer correctly, and the difficulty level of this item is 0,75 included in the easy item category.

4.1.2 Discrimination Power Level

The discrimination power of items is analyzed. Firstly, the researcher made a table according to each student's answer sheets. According to the resulting score, the students were divided into two groups: the lower and upper groups. Then, the researcher calculated the level of discrimination power by using the following formula:

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B}$$

Where:

D: Item Discrimination (Discrimination Power)

BA: Number of top test takers that have the correct answer

BB: Number of bottom test takers that have the correct answer

JA: total participant of top test-takers

JB: total participant of bottom test-takers

(Brown, 2004: 59)

After the discrimination power level is calculated, the researcher used the following classified results from Arikunto's book to indicate the discrimination power level; the item is considered poor, satisfactory, good, and excellent.

Table 4.3.The Classification of item discrimination power

Discrimination Index	Quality
0.0 – 0.20	Poor
0.21 – 0.40	Satisfactory
0.41 – 0.70	Good
0.71 – 1.00	Excellent

(SuharsimiArikunto, 2016 p. 232)

To read the information easier of discrimination power level, each multiple-choice items is interpreted into a classification table as follows:

Table 4.4.The Data Interpretation of Item Discrimination Level

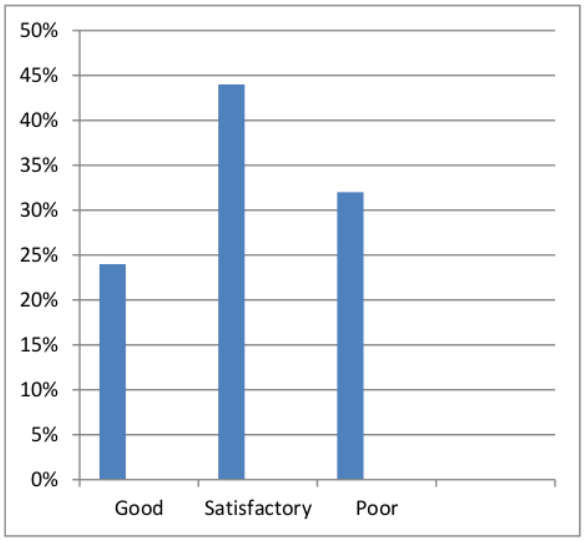
Item Number	**Item Discrimination Index**	**Classification**
1	0.13	Poor
2	0.13	Poor
3	0.50	Good
4	0.38	Satisfactory
5	0.38	Satisfactory
6	0.13	Poor
7	0.50	Good
8	-0.38	Satisfactory
9	0.38	Satisfactory
10	0.25	Satisfactory
11	0.25	Satisfactory
12	0.25	Satisfactory
13	0.50	Good
14	0.25	Satisfactory
15	0.13	Poor
16	0.38	Satisfactory
17	0.50	Good
18	0.63	Good
19	-0.25	Satisfactory
20	0.50	Good
21	0.25	Satisfactory
22	0.00	Poor
23	0.13	Poor
24	-0.13	Poor
25	0.00	Poor

As seen from Table 4.4 above, there were three columns in the table: the first column showed the number of items. The second column showed

the results of the discrimination power level analysis of items. The third column showed about Classification of the discrimination power level. The following the description of the statistically calculated data:

Nama	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	total
Putri Ainun hikmah	1	0	1	0	1	1	1	0	1	1	1	0	1	1	1	1	1	1	0	1	0	1	1	0	1	18
Muzakir	1	0	1	1	0	1	0	0	0	1	0	1	1	0	1	1	1	1	0	1	0	1	0	0	1	14
Lira Firna	1	1	1	1	1	0	1	0	1	0	1	0	1	0	0	0	1	1	0	1	1	0	0	0	1	14
Siti Kirani Fadilla	0	1	1	1	1	0	1	0	0	0	1	0	0	1	1	0	1	1	0	1	0	1	0	0	1	13
Tita Anjalita	0	1	1	0	1	0	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	0	1	13
Asrul Rahman	1	1	0	0	1	0	1	1	0	0	0	1	1	0	0	1	1	0	1	0	1	1	0	0	0	12
Sukmayati	0	1	0	0	1	1	1	0	0	1	1	1	1	0	0	1	0	0	0	1	0	1	0	1	0	12
Raisya Madani	1	1	0	1	1	0	1	0	1	1	0	0	1	0	0	0	1	0	1	0	0	1	0	0	1	12
Kelompok atas	0.63	0.75	0.63	0.50	0.88	0.38	0.88	0.25	0.38	0.50	0.50	0.38	0.88	0.38	0.50	0.63	0.88	0.63	0.25	0.63	0.25	0.75	0.25	0.13	0.75	
Nurhaila	0	1	0	0	1	0	1	1	0	0	1	1	0	0	0	1	0	0	1	0	0	1	0	1	0	10
Sisa Anggriani	1	1	0	0	1	0	0	1	0	0	1	0	1	0	1	0	0	0	1	0	0	1	0	0	1	10
Nurmajidah	0	0	1	0	0	1	1	1	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	1	1	9
Lydia	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	1	8
Ainun Dwi Raham	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1	1	0	1	7
Haikal	0	1	0	0	0	1	1	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	1	7
Emilia	1	0	0	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	6
Nurul Fatiah	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0	5
Kelompok bawah	0.50	0.63	0.13	0.13	0.50	0.25	0.38	0.63	0.00	0.25	0.25	0.13	0.38	0.13	0.38	0.25	0.38	0.00	0.50	0.13	0.00	0.75	0.13	0.25	0.75	
Daya Beda	0.13	0.13	0.50	0.38	0.38	0.13	0.50	-0.38	0.38	0.25	0.25	0.25	0.50	0.25	0.13	0.38	0.50	0.63	-0.25	0.50	0.25	0.00	0.13	-0.13	0.00	

From table 4.4 above, it could be seen that the poor categorized there were 8 (32%) items that are number 1, 2, 6, 15, 22, 23, 24, and 25. The satisfactory items were 11 (44%) items that are number 4, 5, 8, 9, 10, 11, 12, 14, 16, 19, and 21.then The last, the good items were 6 (24%) items that are number 3, 7, 13, 17, 18, and 20. For other information easier to read, the chart below is the percentage of item discrimination levels. Of the 25 items, there were 6 (24%) items that belong to the good criteria, then there were 11 (44%) items classified as satisfactory, and there were 8 (32%) items that included poor categorization.



For more detailed information, the researcher describes each item as follows:

1. Item number 1 has the discrimination power value of 0.13, and this item is categorized as poor discrimination power.
2. Item number 2 has the discrimination power value of 0.13, and this item is categorized as poor discrimination power.
3. Item number 3 has the discrimination power value of 0.50, and this item is categorized as having good discrimination power.
4. Number 4 has the discrimination power value of 0.38, and this item is categorized as having satisfactory discrimination power.
5. Item number 5 has the discrimination power value of 0.38, andthis item is categorized as satisfactory discrimination power.
6. Item number 6 has the discrimination power value of 0.13, and this item is categorized as having poor discrimination power.

7. Item number 7 has the discrimination power value of 0.50, andthis item iscategorized as havinggood discrimination power.
8. Item number 8 has the discrimination power value of -0.38, and this item is categorized as having satisfactory discrimination power.
9. Item number 9 has the discrimination power value of 0.38 and the categorized as having satisfactory discrimination power.
10. Item number 10 has the discrimination power value of 0.25 and the categorized as having satisfactory discrimination power.
11. Item number 11 has the discrimination power value of 0.25 andthe categorized as having satisfactory discrimination power.
12. Item number 12 has the discrimination power value of 0.25 and the categorized as having satisfactory discrimination power.
13. Item number 13 has the discrimination power value of 0.50 and the categorized as having good discrimination power.
14. Item number 14 has the discrimination power value of 0.25 andthe categorized as having satisfactory discrimination power.
15. Item number 15 has the discrimination power value of 0.13 and the categorized as having poor discrimination power.
16. Item number 16 has the discrimination power value of 0.38 andthe categorized as having satisfactory discrimination power.
17. Item number 17 has the discrimination power value of 0.50 and the categorized as having good discrimination power.

18. Item number 18 has the discrimination power value of 0.63 and the categorized as having good discrimination power.
19. Item number 19 has the discrimination power value of -0.25 andthe categorized as having satisfactory discrimination power.
20. Item number 20 has the discrimination power value of 0.50 and the categorized as having good discrimination power.
21. Item number 21 has the discrimination power value of 0.25 and the categorized as having satisfactory discrimination power.
22. Item number 22 has the discrimination power value of 0.00 and the categorized as having poor discrimination power.
23. Item number 23 has the discrimination power value of 0.13 and the categorized as having poor discrimination power.
24. Item number 24 has the discrimination power value of 0.13 and the categorized as having poor discrimination power.
25. Item number 25 has the discrimination power value of 0.00 and the categorized as having poor discrimination power.

4.2 Discussion

The result of the analysis of difficulty level of the multiple-choice items for the English final semester test of tenth grade MA AL-Muthmainnah, six items have to the difficult category, namely item numbers 9, 12, 14, 21, 23, and 24, and this is because the difficulty index of these items in the range 0.00 to 0.30. For example, the items number 12 is as follow:

Figure 4.1

The example of final semester test item no.12

Illegal Medicine Destroyed

Hundreds of packages of medicine and traditional herbs of various brands were confiscated and destroyed by Yogyakarta's Food and Drugs Control officers during its joint operation with the City's Health officers starting on Monday.

The joint operation was conducted by seven personnel at several shops and drugs stores in several parts of the city. "In the fight againts unregistered medicine we'll continue the raids,"said the head of Yogyakarta City's Health Officer, Chairul Anwar on Tuesday. Shop owners found setting the registered medicine would first be warned as they probably did not know they were selling illegal products.

Officers

"The joint operation was conducted by sevenpersonnels at several shops and drugs stores in several parts of the city"

The similar meaning of the underlined word is …

a. Beverages
b. Foods
c. Medicines
d. Tablets
e. Pills

The item included in the moderate categorized, there were 17, namely item numbers 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, 18, 19, and 20, this is because of the difficulty index these items in the range 0.31 to 0.70. For example, the items number 3 is as follow:

figure 4.2

The example of final semester test item no.3

Dear Hafidz,
Please accept our sympathies on your
daughter's death, Naila. Our thoughts and
prayers are with you during this difficult time.
Please know that we are here for you.

With love
Reza, Kiki, and Akbar.

1. Please accept our sympathies on your daughter's death, Naila.
   The antonym of the underlined word is ….
   a. Receive
   b. Reject
   c. Give
   d. Bring
   e. Get

T

The last, there were two items included in easy categorization item numbers 22 and 25. It is because of the difficulty index these items in the range of 0.71 to 1.00. For example, the items number 22 is as follow:

Figure 4.3

The example of final semester test item no.12

Last week I took my five-year old son, Willy, to a musical instrument store in my hometown wanted to buy him a set of junior drum because his drum teacher advised me to buy him one. Willy likes listening to music very much. He also likes asking me everything he wants to know. Even his questions sometimes seem precious for a boy of his age. He is very inquisitive.

We went there by car. On the way, we saw a policeman standing near a traffic light regulating the want passing cars and other vehicles. He blew his whistle now and then. Seeing the policeman blowing his whistle, Willy asked me at once, "Dad, why is the policeman using a whistle, not a drum?"Hearing his unexpected question I answered reluctantly, "Because he is not Phil Collins!"

22. He also likes asking me everything he wants to know.
The underlined word refers to…

a. Father
b. Willy
c. Policeman
d. Phillcollins
e. Drum teacher

In the final semester, test items for the tenth grade of MA AL-Muthmainnah made by the English teacher of the 25 items are categorized into items that have the moderate difficulty level (68%), the difficult level (24%) and easy difficulty level (8%).

The analyzes result of the discrimination power level of the multiple-choice items for the English final semester test of tenth grade MA AL-Muthmainnah made by an English teacher from the 25 items, were 6 (24%) items that belong to the good criteria. There were 11 (44%) items classified as satisfactory. There were 8 (32%) items that included poor categorization. It can be concluded that the items made by English teacher MA AL-Muthmainnah for the final semester test tenth grade have satisfactory

discrimination power (44%). While the poor discrimination power (32%)and the good discrimination power (24%).

From the explanation above, it is necessary to remember a relationship between the two, namely difficulty level and discrimination power. Analysis of the level of difficulty can be used as an indicator to determine the differences between students. Some of the experts believe that the items considered good are the items that have a medium difficulty level. It means that the items have a difficulty index between 0.30 to 0.70. it does not mean that the test items are too easy or too difficult can not be used. It depends on the purpose of the test. Meanwhile, items that have a discrimination index of 0.21 or even higher are recommended for use. Test items that have negative discrimination denote several errors, such as errors in grading the test items, the bottom group of students guessing the answer correctly, or the upper group the student chooses the wrong answer.

CHAPTER V

CONCLUSION AND SUGGESTION

In this chapter, the researcher discusses the conclusion and suggestions from the result of this study

5.1 Conclusion

Based on the results of the analysis of multiple-choice items of English final semester test items for tenth grade made by the English teacher MA AL-Muthmainnah in the academic year 2020/2021, which consists of the difficulty level and discrimination power above, it can be concluded that: the difficulty level of items made by the English teacher for the English final semester test of tenth grade MA AL-Muthmainnah is categorized as having the good difficulty level because the difficulty level of the items has moderate difficulty level (68%), It means the items are not too difficult or too easy.

Meanwhile, the English teacher's discrimination power for the English final semester test of tenth grade MA AL-Muthmainnah was categorized as having satisfactory discrimination power level because the items have satisfactory discrimination power (44%). Thus, discrimination power is

categorized as good discrimination with a low value of (24%) and is higher than the discrimination power level categorized as poor (32%).

5.2 Suggestion

Based on the result of the research, the analysis of multiple-choice items of English final semester test made by the English teacher for tenth grade MA AL-Muthmainnah in the academic year 2020/2021. There are some suggestions addressed to test make and four further research:

1. For teachers, the result of this study can be considered in various policies in improving the quality of the final semester test items and reviewing the items that are not good if used in the following final semester test. Although the items still have are only as gauges students' abilities, but the items made by the teacher should have good quality. Because this test is not only a ranking selection, but the test is also able to provide information or suggestions for test users or students to take what steps in the future should be done.
2. Test items categorized as very good, good, and satisfying can be saved for reference or the next final semester. In contrast, those that are not good and bad should be revised according to the indicators of the causes of failure to become a good problem.
3. Items with a negative discrimination index need to be discarded because they do not differentiate between upper and lower group students.
4. For further researchers, it is advisable to use more samples to obtain more objectives data results.

AN ANALYSIS OF MULTIPLE CHOICE ITEMS OF ENGLISH FINAL SEMESTER TEST MADE BY ENGLISH TEACHER (Study at the Tenth Grade MA AL-Muthmainnah in Academic Year 2020/2021)

ORIGINALITY REPORT

45%
SIMILARITY INDEX

PRIMARY SOURCES

1	repositori.uin-alauddin.ac.id Internet	1439 words —	16%
2	repository.uinjkt.ac.id Internet	869 words —	9%
3	digilib.uinsby.ac.id Internet	423 words —	5%
4	media.neliti.com Internet	236 words —	3%
5	pt.scribd.com Internet	158 words —	2%
6	eprints.iain-surakarta.ac.id Internet	147 words —	2%
7	eprints.undip.ac.id Internet	141 words —	2%
8	jurnal.umj.ac.id Internet	116 words —	1%

9	sukatnowonogiribelajar.blogspot.com Internet	102 words —	1%
10	repository.ummat.ac.id Internet	62 words —	1%
11	repository.iainpalopo.ac.id Internet	52 words —	1%
12	e-repository.perpus.iainsalatiga.ac.id Internet	47 words —	1%
13	etheses.iainponorogo.ac.id Internet	47 words —	1%
14	repo.unhi.ac.id Internet	29 words —	< 1%
15	id.123dok.com Internet	28 words —	< 1%
16	ejournal.kopertais4.or.id Internet	27 words —	< 1%
17	repository.ub.ac.id Internet	18 words —	< 1%
18	repository.uin-suska.ac.id Internet	18 words —	< 1%
19	chonail.blogspot.com Internet	17 words —	< 1%
20	eprints.umg.ac.id Internet	17 words —	< 1%
	www.klis.pf.ukf.sk		

21	Internet	13 words — < 1%
22	lib.unnes.ac.id Internet	10 words — < 1%
23	repository.uinsu.ac.id Internet	10 words — < 1%
24	repository.upstegal.ac.id Internet	10 words — < 1%
25	digilib.iain-palangkaraya.ac.id Internet	9 words — < 1%
26	repo.iain-tulungagung.ac.id Internet	9 words — < 1%
27	view.etextsource.com Internet	9 words — < 1%
28	docobook.com Internet	8 words — < 1%
29	e-journal.metrouniv.ac.id Internet	8 words — < 1%
30	www.bunyanie.com Internet	8 words — < 1%
31	Rohmatul Jannah, Didin Nuruddin Hidayat, Nida Husna, Imam Khasbani. "An item analysis on multiple-choice questions: a case of a junior high school English try-out test in Indonesia", Leksika: Jurnal Bahasa, Sastra dan Pengajarannya, 2021 Crossref	7 words — < 1%

EXCLUDE QUOTES	OFF	EXCLUDE MATCHES	OFF
EXCLUDE BIBLIOGRAPHY	OFF		