

**AN ANALYSIS ON THE ITEMS DIFFICULTY LEVEL OF ENGLISH
SEMESTER TEST AT THE TENTH GRADE STUDENTS OF SMAN 3
JAMBI CITY ACADEMIC YEAR 2016/2017**

A RESEARCH PROPOSAL

**Presented as a Partial Fulfillment of the Requirements for the Attainment of
the Degree of *Sarjana Pendidikan* in English Educational Study Program**



**ENGLISH EDUCATION STUDY PROGRAM FACULTY OF
TEACHERS TRAINING AND EDUCATIONAL
BATANGHARI UNIVERSITY**

JAMBI

2017

APPROVAL

This thesis “AN ANALYSIS ON THE ITEMS DIFFICULTY LEVEL OF ENGLISH SEMESTER TEST AT THE TENTH GRADE STUDENTS OF SMAN 3 JAMBI CITY ACADEMIC YEAR 2016/2017”. Written by Adelia Damayanti, students number 1300888203043 has been corrected and approved to be examined in front of the team of examiners.

Jambi, August 2017

Approved and Accepted by

The First Advisor,

The Second Advisor

Dra. Hj. Wennyta, M.Pd.

Khidayatul Munawwaroh, M.Pd.

The Head of English Educational
Training

The Dean of Teacher

Study Program

and Education Faculty

Dra. Hj. Wennyta, M.Pd.

Abdoel Gafar, S.Pd., M.Pd.

LETTER OF RATIFICATIONS

Name : Adelia Damayanti

Students Number : 1300888203043

Defending her thesis “**An Analysis on the Items Difficulty Level of English Semester Test at the Tenth Grade Students of Sman 3 Jambi City Academic Year 2016/2017**”, was examined and stated to pass the examination on:

Day : Friday

Date : August 25th 2017

Team of Examiners

NAMES

POSITION

SIGNATURE

Dra. Hj. Wennyta, M.Pd.

First Advisor

Khidayatul Munawwaroh, M.Pd.

Second Advisor

Dr. Sri Marmoah, M.Pd

First Examiner

Nurul Fitri, S.S., M.Hum

Second Examiner

Ratified by

**The Head of English Educational
Training**

Study Program

The Dean of Teacher

and Education Faculty

Dra. Hj. Wennyta, M.Pd.

Abdoel Gafar, S.Pd., M.Pd.

STATEMENTS OF WORK'S ORIGINALITY

I honestly declare that the thesis I wrote does not contain the works or parts of the works of other people, except those cited in the quotations and bibliographies, as a scientific paper should.

Jambi, August 12th 2017

The Writer

Adelia Damayanti



ACKNOWLEDGEMENT

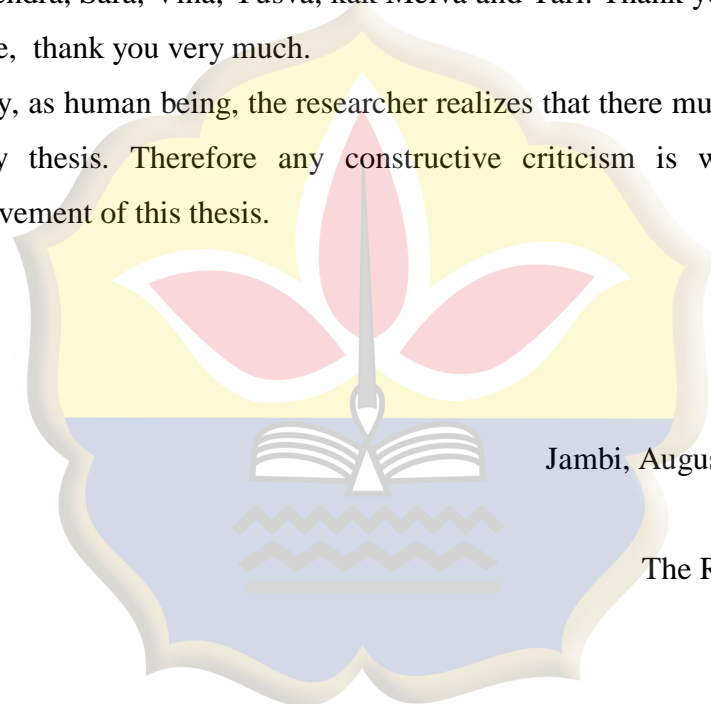
Praise and pray that ALLAH SWT has given the gift of his grace and the researcher can finish a thesis entitled “An Analysis on The Difficulty Level Items of English Semester Test at The Tenth Grade Students of Sman 3 Jambi City in The Academic Year 2016/2017”. Success in preparation of this thesis can not be separated from the help of various parties.

Many people deserve special recognition for their invaluable help it accomplishing this thesis. I own enormous debts of gratitude to several people. So, it is an excellent opportunity to acknowledge my warm, thanks for the following remarkable figure :

1. H. Abdoel Gafar, S.Pd, M.Pd. as Dean of the Faculty of Teacher Training and Education of Batanghari University.
2. Dra.Hj. Wennyta, M.Pd as The Head of English Education Program and also as my first advisor, thank you for the support,time, advice and sincerity so I can finish my research well.
3. The second advisor, Khidayatul Munawwaroh, M. Pd. who has been willing to take the time to guide this thesis and giving the researcher her kindness.
4. All lectures in English Education Program Teacher Training and Education Faculty who have provide the science and education are very significant.
5. The best appreciation is for my lovely mom Askinah as the best mom ever in the world and also my superdad i love u both. Thank you for always praying, encouraging, giving love, giving material i need, working hard, and always trying to do the best for me. Thank you for the love you have given. I can survive and have spirit to face everything until today just because of them.

6. For my dearest and the special one Ari Rosadi, thank you so much. Thank you for always supporting and accompanying me whenever i need. Thank you for the advices you gave to me to finish this thesis. Thank you for the great love you have given.
7. Big family at SMA N 3 Jambi City who helped the researcher to complete the data needed to finish this research.
8. All of my friends in English Educational Study Program academic year 2013, especially : Intan, Nina, Nova, Longga, Walsa, Reski, Trik, Can, Hendra, Sara, Vina, Yusva, kak Melva and Tari. Thank you for supporting me, thank you very much.

Finally, as human being, the researcher realizes that there must be weaknesses in my thesis. Therefore any constructive criticism is welcome for the improvement of this thesis.



Jambi, August 12th, 2017

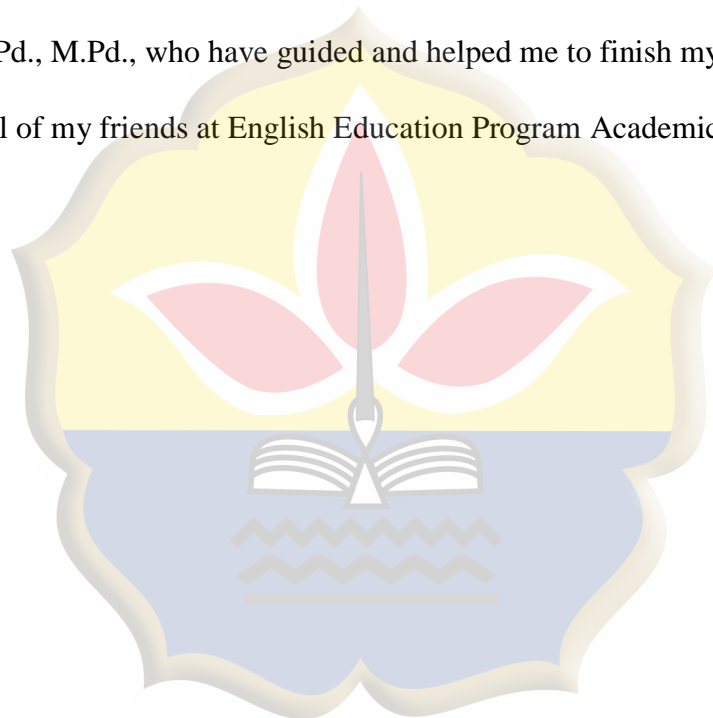
The Researcher

Adelia Damayanti

DEDICATION

This work is dedicated to:

- Allah SWT who always give His graces and blessing.
- My great parents, Suhara and Askinah, who always give me support, advices, motivation, attention, and pray for me, my utmost gratitude for their affection and endless love.
- My Advisors, Dra. Hj. Wennyta, M.Pd. and Khidayatul Munawwaroh, S.Pd., M.Pd., who have guided and helped me to finish my thesis.
- All of my friends at English Education Program Academic Year 2013.



MOTTO



ABSTRACT

Damayanti. Adelia. 2017. *An Analysis on The Difficulty Level Items of English Semester Test at The Tenth Grade Students of Sman 3 Jambi City in The Academic Year 2016/2017*. Sarjana Thesis. English Educational Study Program. Faculty of Teacher Training and Educational Sciences. Batanghari University of Jambi. Advisor I: Dra. Hj. Wennyta, M.Pd. Co-advisor: Khidayatul Munawwaroh, S.Pd, M.Pd.

Keywords: English Semester Test, Item Analysis, Difficulty Level.

The purpose of this study is to measure the difficulty level of the English semester test items for the tenth grade students at SMA N 3 Jambi City in the academic year 2016/2017. The sample is 40 students of tenth grade science two and taken by random sampling with lottery. This study is conducted to analyze the test items whether it is conclude in easy, moderate, or difficult items. This study is categorized in mix method research, those are quantitative and qualitative. The instrument used in this study is documentation. The researcher take 40 students' answer sheets as the data that will be counted how many responden can answer of each item and has been counted by using formula. The research find that there are 35 items in easy level, 2 items in moderate level, and 3 items in difficult items. And all the items can be said in easy level with the result 0,85 index of difficulty. The researcher also find that there are 12 valid items, and 28 invalid items.

TABLE OF CONTENTS

Cover.....	i
Approval	ii
Letter of Ratification.....	iii
Statements of Work’s Originality	iv
Acknowledgement	v
Dedication.....	vi
Motto	vii
Abstract	viii
Table of Contents.....	ix
List of Tables	x
List of Appendices	xi

CHAPTER I : INTRODUCTION

1.1 Background of the Research.....	1
1.2 Formulation of the Research.....	4
1.3 Limitation of the Research.....	4
1.4 Purpose of the Research.....	4
1.5 Benefit of the Research.....	5
1.5.1 Theoretical Benefit.....	5

1.5.2 Practice Benefit	5
1.6 Definition of Key Terms..	6
CHAPTER II : REVIEW OF THE RELATED LITERATURE	
2.1 Theories of Evaluation	7
2.1.1 Principle of Evaluation.....	8
2.1.2 Objectives of Evaluation	10
2.1.3 Steps of Evaluation.....	11
2.2 Test as Evaluation Instrument	12
2.2.1 Definition of the Test	12
2.2.2 Types of Tests	13
2.2.2.1 Function of The Test.....	13
2.2.2.2 Way of Scoring	15
2.2.3 Item Analysis.....	19
2.2.4 Difficulty Level of the Item.....	20
2.2.5 The Importance of Item Analysis	22
2.3 Previous Studies	23
2.4 Conceptual Framework	26
CHAPTER III : RESEARCH METHODOLOGY.....	29
3.1 Research Design	29
3.2 Population.....	29
3.3 Sample	30
3.4 Technique of Collecting Data.....	30
3.5 Technique of Data Analysis	31
CHAPTER IV : FINDING AND DISCUSSIONS.....	33
4.1 Findings	33

4.2 Discussion.....	35
CHAPTER V : CONCLUSION AND SUGGESTION.....	43
5.1 Conclusion.....	43
5.2 Suggestion	43
REFERENCES.....	45
APPENDICES	48



LIST OF TABLES

Table 2.1 Classification of Difficulty Level	22
Table 3.1 Population of the Research	30
Tabel 3.2 Classification of Difficulty level by Arikunto	32
Table 4.1 Result of Item Classification.....	33
Table 4.2 Category of difficulty level of the English semester test items	41



LIST OF APPENDICES

APPENDIX 1 RESEARCH DATA TABULATION	48
APPENDIX 2 VALIDITY ITEM RESULT	50
APPENDIX 3 R-TABLE	52
APPENDIX 4 QUESTION PAPER	53
APPENDIX 5 STUDENTS' ANSWER SHEETS	62
APPENDIX 6 ANSWER KEY	63



CHAPTER I

INTRODUCTION

1.1 Background of The Research

In Indonesia, English is the most used foreign language, because English in important position, it is taught from elementary school as an alternative subject and from the first grade of junior high school up to third grade of senior high school as one of the compulsory subjects. The students of junior high school and senior high school are hoped to have enough competence in English because it becomes a subject tested in UN (Ujian Nasional). The government always make an effort to increase the quality of education in general and English teaching in particular.

In the teaching learning process teachers have to make learners different at the end of an instructional unit from what they are before. Learners who have completed a unit of learning tasks should be different from this who have not done it. To assess learner's achievement it is useful for teacher to conduct a test or examination. Test or evaluation is an important part of every teaching and learning experience. Evaluation is an important part of every teaching and learning experiences. It gives big contribution for the teaching and it provides an information about the students' progress which can be used by the teachers to manage the learning task and students. As stted by Rea and Germain (1992:3) “ Evaluation is important for the teacher because it provides a wealth of information yto use for the future direction of classroom practice, for the planning of courses and for the management of learning tasks and students. Bachman(1990:22) also

states that “ Evaluation can be defined as the systematic gathering of information for the purpose of making decision.

Depending upon the decision being made and the information a teacher needs in order to inform that decision, testing often contribute to the process as the implementation to collect data. Nitko(1983:6) states “ A test is defined as a systematic procedure for observing and describing one or more characteristic of a person with the aid of either a numerical scale or category system. According to Henning(1987:1) “ Tests in general is to pin point strengths and weakness in the learned abilities of students”. Teachers need to do the test because through the test they are able to find out the students’ achievement in mastering the lessons that have been taught and to evaluate the effectiveness of a new teaching method used and the teaching material. Supported by Valette(1977:5) “...through tests the teacher can evaluate the effectiveness of a new teaching method, of a different approach to a difficult pattern, or of a new material”. To measure the students’ learning progress at school, a teacher commonly administers two kinds of test ; formative and summative test. the former test is held earlier than latter test which is held at the end of semester. Through both tests, a teacher can measure the students’ achievement level and the degree of how far the instructional objectives of learnig be eaccomplished by them.

Good English test also help students to learn the language by requiring them to study hard, emphasizing course objectives, and showing them where they need to improve. For the teachers, the test plays several important roles. Such as, to provide insight into ways of improving the evaluation process and to provide means of diagnosing their own efforts if they taught effectively.

For getting accurate measures a test must have a good quality, because a good test doesn't only influence the student's learning, but also influences the teacher to improve teaching and learning process. Heaton (1998:13) supports that "Test may be constructed primarily as a device to reinforce learning and to motivate the student's performance in language".

As the accuracy of a test result influences the motivation of student learning, so the test administered must reflect a good test. Heaton (1998:152-156) states that "a good test is a test which has the criteria of validity, reliability, and practicality. Besides that, it must have discriminating power and difficulty level". A test can be valid if the test can measure what is supposed to measure. It can be reliable if the result of the test is the same even though the test administered to the same level students in the next time. And it can be practical if it's easy to do administer.

The matter, which is often forgotten by the teacher is the follow up of the test implementation pertaining to the test item itself. In fact, they do not criticize whether or not all items have fulfilled the criteria above. Therefore, it really required an analysis of the test items, that is namely item analysis. Through analyzing test item teacher can identify good item and the poor item and differentiate between student who have done well and poorly.

The quality of the test can be seen from the point because of the problem analysis. Activities of analyzing the item is an activity that the teacher must do to improve the quality of the problems created. The goal is to study and examine each item in order to obtain a quality question before the question used or after use. In addition, the purpose of the item analysis is also to help improve the test

through revision or discard the problem that is not effective, as well as to know the diagnostic information on the students whether they have or have not understood the material being taught.

Based on the researcher observation during teaching practice in SMA N 3 Jambi City, the researcher found that first, when the students doing the evaluation, most of them failed in the evaluation. Second, It makes the researcher curious to know the problem. The researcher wants to research on the items for the examination. Because success or not a test is depend on the quality of question.

Based on the background above, the researcher wants to conduct research on *“An Analysis on The Difficulty Level Items of English Semester Test at The Tenth Grade Students of Sman 3 Jambi City in The Academic Year 2016/2017”*.

1.2 Formulation of The Problem

Based on the background of the problem that has been stated above, the issues discussed in this study *“How are the difficulty level items of the English semester test of the tenth grade student at SMAN 3 Jambi City?”*

1.3 Limitation of the Problem

The researcher limits the research and focus only on the difficulty level of the English semester test of the tenth grade students at SMAN 3 Jambi city to find out the difficulty level from easy, moderate, and difficult item of English semester test.

1.4 Purpose of the Research

Referring to problem mentioned in the formulation of the problem above, this research aims to know the difficulty level items of the English semester test of the tenth grade student at SMAN 3 Jambi city.

1.5 Benefit of the Research

This research is conducted with the hope of providing benefits to as follows:

1. Theoretical Benefit

The results of this study are expected to increase scientific insight on the theory of analysis about the right evaluation, and provide an overview which is really about the quality of the grain matter of the end of tenth grade students as an effort to improve the quality of the questions used.

2. Practical Benefit

a. For Teacher

To know the quality of the question used as feedback evaluation for improvement or improvement in teaching especially in the preparation and making of learning evaluation questions.

b. For Students

To measure of the success rate of students in learning can be more accurate.

c. For Researcher

This research is expected to be used as material for do more research. Moreover, it can be used for do the development of item analysis process on others subject.

d. For Reader

This research can be used as benchmarks and evaluation materials for the team making the matter of order can further improve the quality of the created questions, resulting in results.

1.6 Definition of Key Terms

A. Test

Test is said to be reliable if the test has been given repeatedly and provide consistent results (Arikunto, 2013: 74). If there is no subjectivity factor in the scoring system, the test can be said as objective. Tests are 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 labor, and a long time (Arikunto, 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.

B. Item Analysis

Item analysis is a process in which responses to test items are examined in order to test the quality of those items

C. Difficulty Level

Level of difficulty is the extent to which an item is easy or difficult for the proposed group of test-takers (Brown, 2004:58). Grondlund (1993:103) states that difficulty level of an item in a test is the percentage of students who answer test items correctly.

CHAPTER II

REVIEW OF THE RELATED LITERATURE

2.1 Theories of Evaluation

According to Arikunto (2013; 3) evaluation is a process of data collection to determine how far, in what way, and how the goal of education has been achieved. According to Sudijono (2011; 1) evaluation in education is an activity or process determination of the value of education, so the quality or the result can be known. According to Zainal Arifin (2011: 5) “evaluation is a process systematically and continuity to determine quantity (meaning and value) from something, based on consideration and certain criteria for decision making”. According to Djemari Mardapi (2007: 8) “evaluation is combination of activities to improve quantity, performance, or productivity institution in implementing the program”.

Ralph Tyler on Arikunto (2009: 3), “evaluation is a process of collecting the data to determine how far the purpose of education has been achieved”. The definition broader stated by Cornbach and Stufflebeam on Arikunto (2009: 3) “evaluation is not about process purpose measuring process, but it is also used to make a decision”. Both these experts, consider to define evaluation not only to assess the results of study but the evaluation also as a factor in the decision making.

So, the researcher can draw that evaluation is a process for measure the learning outcomes of learners from the beginning of the process to the end. Data is taken during the learning process is the end. The learning data is collected to be analyzed to know how far the goal of learning is achieved.

2.1.1 Principle of Evaluation

According to Arifin (2011: 30-31) to obtain a good evaluation, evaluation should have to refer on general principles, as follows:

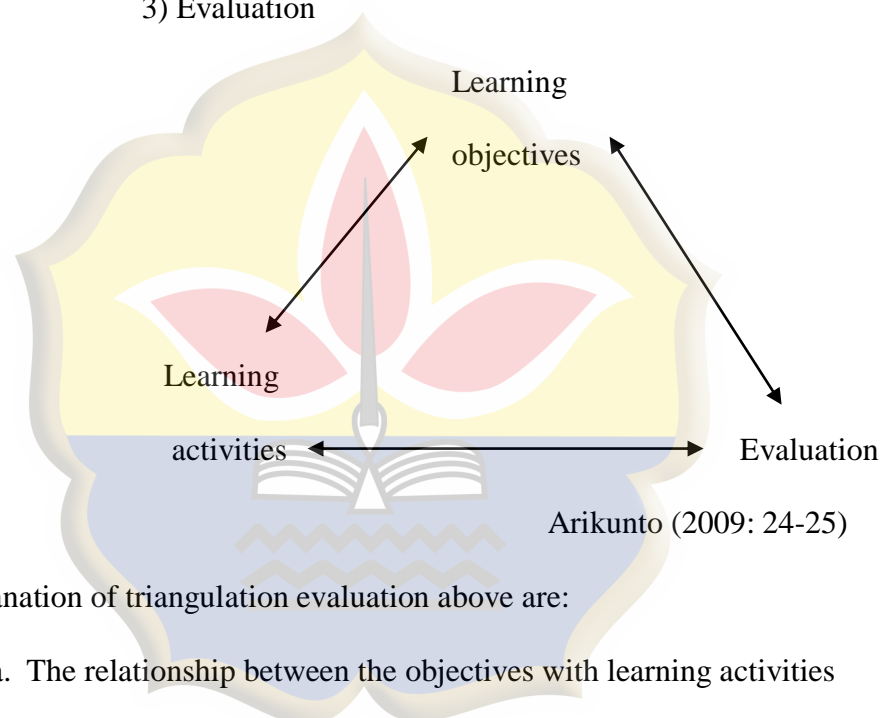
- a. Continuity, evaluation should not be done in learning own incidentally because it is a process continuous. Therefore evaluation should be done continuously. The evaluation results obtained at a time should always connected with the result. The students do not study can be seen just from a product but also a process in fact the input.
- b. Comprehensive, in evaluating an object, the teacher should take all the objects as the evaluation. For example, if the object of evaluation is students, then all aspects of the students, should be evaluated both involving cognitive, an affective, or psychomotor aspects.
- c. Fair and Objective, in implementing an evaluation, teachers should do fairness. "Fair and objective" is easy to say, but it is difficult to do. All student have same equality without discrimination. The teacher also should be objective one. Therefore, all the bad be have like and dislike, desires, and prejudice negative should be removed. Evaluation should be based on reality, no manipulation.
- d. The cooperative, in implementing evaluation the teacher should make a good cooperation with all parties, such as parent of student, the other teachers, headmaster, including with the students

themselves. It is done to make all parties satisfy with the evaluation result.

- e. Practically, easy to use by all parties. If the instrument to filling the pre requirement but difficult to use, so the instrument not practice.

There is one general principle namely triangulation, they are:

- 1) Learning objectives
- 2) Learning activities
- 3) Evaluation



The explanation of triangulation evaluation above are:

- a. The relationship between the objectives with learning activities
Learning activities that design in the lesson plan in line with the objectives of learning, so every activities of students in the class is aimed to achieve the objective of learning.
- b. The relationship between the objectives with evaluation
Evaluation is the activity of collecting data to measure the goals that have been achieved. Then, in arranging instrument and techniques for evaluation should refer to the objectives that have been formulated.

- c. The relationship between the evaluation with learning activities

Evaluation should refer or adapted with activities learning. For example, in teaching and learning activities the teacher more oriented in skill, then evaluation also make to measure the aspect skill of students.

2.1.2 Objectives of Evaluation

Purpose of teacher in doing evaluation is to know how far a student reach of learning. Besides that, the result of evaluation will give an interpretation on the ability and the level of knowledge the student with the materials that delivered by teachers.

According to Arifin (2013:14), the purpose of evaluation learning is to know the effectively and efficiency of the learning system, which concerns about the objectivities, material, method, media, a source of learning, environment and assessment system. According to Sudijono (2011: 16-17), the objective of evaluation in education into the general objectivities and a special objectivities.

1) General objectivities

Generally, the objectives of evaluation in education field are:

- a. To obtain evidence data, which will be used clue of ability and success level of the student.
- b. To see the effectiveness of the learning that methods used by teacher.

2) Specific objectivities

The evaluation being a special purpose of activity in education are:

- a. To stimulate the activity of student in education program.

- b. To search and discover cause failure and successes of factors in following education program.

2.1.3 Steps of Evaluation

Sudijono (2011: 59) explains the steps of evaluation, as follows:

1. Arrange a plan evaluation of learning outcomes. Planning evaluation of learning outcomes include six types of activities, namely
 - a) Formulating the purpose of implementation of the evaluation.
 - b) Establish the aspects that will be evaluated, such as cognitive, affective aspects, or psychomotor aspects.
 - c) Select and determine the techniques that will be used in the evaluation.
 - d) Arrange tools that will be used in the measurement and assessment of student learning outcomes.
 - e) Determine benchmarks, norms or criteria that will be used as a handle or a benchmark in providing interpretations of the evaluation data.
 - f) Determine the frequency of the activity evaluation of learning outcomes.
2. Collecting the data. Collecting the data is to carry out the measurement or observation, interviews or questionnaires with the use of certain instruments.

3. To verify the data. Verification of data is intended to be able to separate good data (data that will be able to clarify the picture to be obtained regarding the individual or group being evaluated) of the data less well (data that will blur the image that would be obtained if the data is joined processed).
4. Process and analyze the data. In processing and analyzing the results of the evaluation carried out with a view to giving meaning to data that have been collected in the evaluation. In processing and analyzing evaluation data can be used statistical techniques or techniques no statistic, depending on the type of data to be processed and analyzed.
5. Provide interpretations and giving conclusions. Interpretation the data evaluation is a verbalization of meaning contained in the data that has processing and analyzing.
6. Follow-up results of the evaluation. Data evaluation results are compiled, organized, processed, analyzed and summarized to known the meaning contained so that evaluators will be able to take decisions or formulate policies that are deemed necessary as a result of the evaluation activities.

2.2 Test as Evaluation Instrument

2.2.1 Definition of the Test

According to Bachman (1990:20) ‘A test is procedure designed to elicit certain behavior from which one can make inferences about certain characteristic of an individual’. While Tinambuan (1988:310) states “A test is a set of

questions, each of which has a correct answer, that examines usually answer orally or in writing. Beside that Airasian (2008:9) said that “A test is a formal, systematic, usually paper-and-pencil procedure for gathering information about pupil’s performance.

So the researcher can draw that test can be instrument, techniques, or procedures to have the students’ respond through tasks or performance in the form of set of questions must be answered in order to achieve the teaching learning objectives. In short, a test is a measurement instrument designed to assess a specific sample of individuals’ behavior.

2.2.2 Types of Tests

There are many kinds of test used to measure students’ achievement that can be used in an evaluation process. The type of test can be classified into two types, namely; function and way of scoring.

2.2.2.1 Function of The Test

According to Harrison (1983:26). The types of functional test can be categorized into four types: placement test, diagnostic test, achievement test, and proficiency test.

a. The Placement Test

Placement test is used to place a student to appropriate level or section of language curriculum or school. It usually happens in the beginning of course. According to Tinambuan (1988:7).

A placement test is designed to determine student performance at the beginning of instruction, thus, it is designed to sort new students into

teaching groups, so that they can start a course at approximately the same level as the other students in the class. It is concerned with the students' present standing, and so relates to general ability rather than specific points of learning. As a rule the results are needed quickly so that the teaching may begin.

b. The Diagnostic Test

According to Lado (1961:361) diagnostic test is designed to diagnose a particular aspect of a language. Diagnostic tests are also achievement tests, but they are characterized by one distinctive feature, namely that they are designed to show specific weaknesses and strengths within the skills or elements measured.

c. Achievement Test

According to Remmers (1990:9) "achievement tests are designed to measure relative accomplishment in specified areas of work. In the other hand, Hughes(2003:13) states that the purpose of achievement test as its name reflects is to establish how successful individual students, groups of students, or the courses themselves have been in achieving objectives. In another point of view Tinambunan (1988:19) said that "the degree purpose of achievement test is designed to indicate degree of students' success in some past learning activities. Namara (2000:7) states "achievement tests relate to the past in that they measure, what language the students have learned as a result of teaching. Based on the argumentation above about achievement test. The researcher can conclude that the achievement tests are intended to measure how

effectively students have mastered the lesson and how far they have reached the instructional objectives.

2.2.2.2 Way of Scoring

Based on manner of scoring, the type of test item is divided into two general types: objective and subjective test. Heaton (1977:25) states that “subjective and objective test are terms used to refer to the scoring of tests”.

a. Objective test

An objective test item is any test item that there is only a single correct answer. In this test, the students must select one option from some alternative. According to Vallete (1997:6) “An objective test item is any item for which there is a single predictable correct answer.

The objective test item commonly used in classroom testing are true-false, multiple choice, matching, and short answer. According to Grondlund (1968:25) said that “these test item include all of the selection-type items-multiple choice, true false, and matching.

1) True-False

True-False is simply a declarative statement which the students must judge as true or false. According to Stanley (1964:17) states that “true-false item is referred to alternative response item; the item asks the students to answer with the “true” if it conforms to the truth or “false” if it essentially incorrect.

Thus, the item provides the students with a choice of two alternatives, so the students have possibility to guess the answer and

sometimes it will be the right answer. In the other word, students indicate whether a statement is true or false.

2) Multiple-choice item

The multiple choice item consist of a stem, which presents a problem situation, and several alternatives, which provide possible solution to the problem. The stem may be a question or an incomplete statement. The alternatives include the correct answer and several plausible wrong answer, called distracters. Their function is to distract those students who are uncertain of the answer. According to Nitco (1983:190) “A multiple-choice item consists of one or more introductory sentences followed by a list of two or more suggested responses from which the examinee chooses one as the correct answer.

Example:

In objective testing, the term objective refers to the method of...

- a. *Identifying the learning outcomes*
- b. *Selecting the test content*
- c. *Presenting the problem*
- d. *Scoring the answer*

3) Matching

The matching test item consists of two parallel columns with each word. Number of symbol in one column is being matched to a word, sentence or phrase in other column. This type of item is employed widely in situation where relationship of more or less similar ideas, facts, and

principles are to be examined or judged, in this type, students indicate relationship between a set of premises and a set of responses.

Example: 1. The drives a car a. Doctor

2.The checks the patience b.driver

This kind of test is an effective way to students’ recognition of the relationships between words, definitions, events, dates, categories, examples, and so on.

b. Subjective Test Item

According to Vallete (1997:10) “subjective item is one that does not have a single right answer”. It means that the scoring is inconsistent and the answer of the question is in form of composition where the students are given a chance to relate their idea or argument in their own words. In the other word, Madsen (1983:8) states that “subjective tests, like translation and essay, have the advantage of measuring language skill naturally, almost the way English used in real life.

The subjective tests that are commonly used in classroom are completion, short-answer, and essay item.

1) Completion

According to Tinambuan (1988:61) the completion item is a written statement that requires the examiner to supply the correct word or short phrase in response to an incomplete sentence, a questions or a word association. Completion test can be used effectively to measure the recall of terms, dates, and names. According to Noll (1965:140) “ the completion item and short answer item are both supply type test items,

but in the short answer type, the blank is nearly always at the end, whereas in the completion, type of the blank may occur everywhere in the statement.

2) Short-answer Item

According to Noll (1965:138) “the short answer item consists of a question. Which can be answered with a word or short phrase. According to Noll (1965:138) “generally, teachers prefer to use the short answer type question, probably because they think it has some advantages. It is relatively easy to construct, it also gives the teacher some opportunity to see how well students can express their thought and it is also no difficult to score or mark than the essay question. However it is difficult to phrase the short answer question, so that only one answer is correct. And this type of question will be more useful only in testing knowledge of facts and quite specific information.

3) Essay Test

As what Gronlund (1985:65) states that:

“Essay tests are inefficient for measuring knowledge outcomes...but they provide a freedom of response which is needed for measuring certain complex outcomes...These include the ability to create...to organize...to intergrate...to express...and similar behaviors that call for the production and synthesis of ideas”.

Finally, from the explanation above about objective test and subjective test concerned on the essay test, the writer concludes that for the measurement of most knowledge outcomes we would use objective test items to take advantage of their

more extensive sampling and greater reliability. For the measurement of such complex learning outcomes as the ability to create, organize, and evaluate ideas, however, the teacher would use essay questions despite their limitation.

2.2.3 Item Analysis

After a test has been administered and scored it is usually desirable to evaluate the effectiveness of the items. This is done by studying the students' responses to each item. When formalized, the procedure is called item analysis. According to Nitko (1983:342) "item analysis refers to the process of collecting, summarizing, and using information about students' responses to items. Meanwhile Madsen (1983:180) states that "The selection of appropriate language items is not enough by itself to ensure a good test. Each question needs to function properly; otherwise, it can weaken the exam. Fortunately, there are some rather some simple statistical ways of checking individual item. This procedure is called 'item analysis'.

According to Arikunto(2009:205) "item analysis is a systematic procedure that will give special information on items that will be arranged. While Hopkins(1998:254) states that "the feedback on individual items can help the instructor to identify points or concepts that are in need of review and further instruction. On the other opinion Tinambunan (1988:137) said that " item analysis usually concentrates three vital features: level of difficulty, discriminating power, and the effectiveness of each alternatives. Thus, item Analysis information can tell us if an item is too difficult or too easy, how well it discriminated between high and low scores on the test, and whether all the alternatives functioned as intended.

2.2.4 Difficulty Level of The Item

In this kind of item analysis, level of difficulty have many other names, these are: item facility, item difficulty, item easiness, p-value, or it can be marked by IF. The difficulty level of item means the percentage of students who answer correctly each test item. According to Nitko(1983:288) “ the item difficulty is fraction of the persons taking an item who answer it correctly.” Heaton(1977:178) states that “ the index of difficulty “(of facility value)of an item simply shows how easy or difficult the particular item provide in the test. The index of difficulty (facility value) is generally expressed as the fraction (percentage) of the students who answered the item correctly”. A good test item should have a certain degree of difficulty. It may not too easy or too difficult because the test that is too easy or too difficult will yield same score distribution that make it hard to identify reliable differences in achievement between the students who have done well and these who have done poorly. According to Bachman(2004:151) “ item difficulty is defined as the proportion of the test takers who answered the item correctly, and the item difficulty index, p, values can be calculated on the basis of test takers response to the item”. As Thoha(2003:145) states that “ the good item is the item which the level of difficulty can be known. Not too easy and not too difficult. Because the level of difficulty has the correlation with the discriminating power. If the item has the maximal difficulty level, so the discriminating power will be low, and so if the item is too easy it will not has the discriminating power”. As Arikunto(2013:222) says that “ the good item is item that is not too easy and not too difficult. The item which is too easy is not stimulate the students’ effort to solve the item. In the other hand the item which is too difficult will caused the

students give up and has no spirit to solve the item and do not want to try again because the item is out of reach”.

From the explanation above, the researcher can draw that difficulty level is important for a test to know which one the difficult and easy item, and also to measure how success the items.

To measure the difficulty level of each item, the researcher uses the Arikunto's formula; the formula is like this:

$$P = \frac{B}{JS}$$

Explanation:

P : Facility value or item of difficulty that we are looking for

B : The total number of person who get the item correct.

JS : The number of students who took a test.

Source Arikunto(2013:223)

After calculating the difficulty of each item, the writer calculates the index of difficulty of all item by Arifins's formula;

$$P = \frac{\sum b}{N}$$

P : difficulty level of all items

B : difficulty level of each items

\sum : sigma (total)

N : total numbers of test items.

To know the criteria of difficulty level of each item and all items, the researcher uses the measurement level referred to Suharsimi Arikunto's book. If the FV is:

Table 2.1 classification of difficulty level

- | |
|--|
| <ul style="list-style-type: none">• Item with P 0,00- 0,30 is difficult item• Item with P 0,30-0,70 is moderate item• Item with P 0,71-1,00 is easy item |
|--|

The level of facility value shows the easiness of difficulty of test item for that group. So, the level of facility value is influenced by the student's competence. The result will be different if the test is given to another group of learners or student.

2.2.5 The Importance of Item Analysis

According to Oller (1979:245) "Item analysis is an important and necessary step in the preparation of good multiple-choice tests: Nitko(1983:284) states that " for teacher made test, the following are among the important uses of item analysis: determining whether an item functions as teacher intends, feedback to the teacher about student difficulties, are for curriculum improvement, revising the item and improving item writing skills.

- a. Determining whether an item function as teacher intends.

Then item will function properly if the test item tested is able to distinguish those who master the learning objectives from those who do not. To differenf between them, the test item should have certain level of difficulty, discriminating power and the effectiveness of distracters. Therefore item analysis should be done.

b. Feedback to students' performance and as a basis for class discussion.

After knowing the students' responds to the item. The students' performance can be known and the students' error can be corrected and the test items that are felt difficult for most of the class can be discussed in their class.

c. Feedback to the teacher about students' difficulties

The result of item analysis will be useful for teachers to know the major types of students' difficulties in learning. So they know material needs to be reviewed in next learning.

d. Area for curriculum

By item analysis, it can be known what kind of items which are felt difficult by students or certain errors occur often, may be the item is not compatible to be taught in a school program. So curriculum may be needed to be revised.

2.3 Previous Studies

Research conducted by Muhammad Taufan Ruspidu (2014) with the title "*Analisis butir soal ujian semester gasal ekonomi akuntansi kelas XI IPA SMA Negeri 11 Yogyakarta Tahun Ajaran 2013/2014.*" The result of this research showed that: (1) valid question are 18 items (45%), invalid question 22 items (55%). (2) reliability whole items in the category of low index is 0,477. (3) Very bad discrimination index is 6 items (15 %), bad items is 5 (12.5 %), enough 16 items (40%), good 11 items (27,5%) and very good 2 items (5%). (4) Based on the level of difficulty 7 question (17,5%) is difficult, intermediate 13 items (32,5%) and easy question is 20 items (50%). (5) Based on distractor efficiency a very good that question is 4 items (10%), good question is 8 items (20%), enough

question is 10 items (25%), bad question is 8 items (20%) and very bad question 10 items (25 %).

Research conducted by Rika Amelia (2010) with the title “An Anlaysia of the English Summative Test Items Items in Terms of Difficulty Level for the Second Year Students of MTs Darul Ma’arif Jakarta”. The research is purposed to measure theh difficulty level of the Eng;ish summative test items by calculation the students’ correct response from the upper and lower group with $J > B$ Heaton’s formula referred from his book “*Dasar-Dasar Evaluasi Pendidikan*” that there are 20 items regarded as difficult item because they are at difficult level, ranges from 0,01 up to 0,30. Twenty one items regarded as good items because they are at moderate level. Ranges from 0,31 up to 0,70. And there are 9 items regarded as easy items because they are at easy level, it ranges from 0,71 up to 1,00. From this information, it can be counted the difficulty level of the items with the total number of students is 0,45. So, it can be said that the english summative test items for the second year students of Mts. Darul Ma’ariff qualified as a good test seen from the difficulty level of all item which is at moderate level, because it ranges from 0,30 up to 0,70.

Khan, Ishrat, and Khan (2015) conducted a research in India entitled “Using Item Analysis on Essay Types Questions Given in Summative Examination of Medical College Students”. The objective of their study is to calculate the Item analysis of questions given in terminal examination of MBBS students and observed the adequacy of questions framing in examination. The study contains 150 medical students undergone terminal examination consists of questions of essay type, structured essay type and short answer questions are

given to the students. They divided them into high ability group and low ability group. Mark range is decided for essay type, structured essay type and short answer type questions and Facility Value (FV) and Discrimination Index (DI) is calculated. The Results of study described that facility values of 62.5% questions are come under recommended and acceptable range and 50% questions are come under acceptable range. Discrimination value of 100% questions are come under recommended and acceptable range where 75% questions are come under acceptable range. The Importance of item analysis is highlighted from their results. For improvement of examination items with average difficulty and high discrimination should be implemented into future examinations to improve test scores.

Boopathiraj and Chellamani (2013) in India conducted a research entitled: “Analysis of Test Items on Difficulty Level and Discrimination Index in the Test for Research in Education”. The objective of their study is to find out the item difficulty and the power of discrimination of Multiple Choice test Items. In this study all student-teachers who are studying Master of Education in Tamilnadu comprise the population of the study. Random sampling is adopted for this research work and 200 students are taken. The sample contituted both male and female student-teachers. It used 60 items for data collection which is developed from the syllabus of Research in Education for Master of Education under Tamilnadu Teachers Education University by the researchers with the help of some subject expert. The results of the research showed that thirteen items out of 60 (21%) are rejected either due to difficulty level or discrimination index. Thirty five items (58%) are accepted without revision while 12 items are accepted

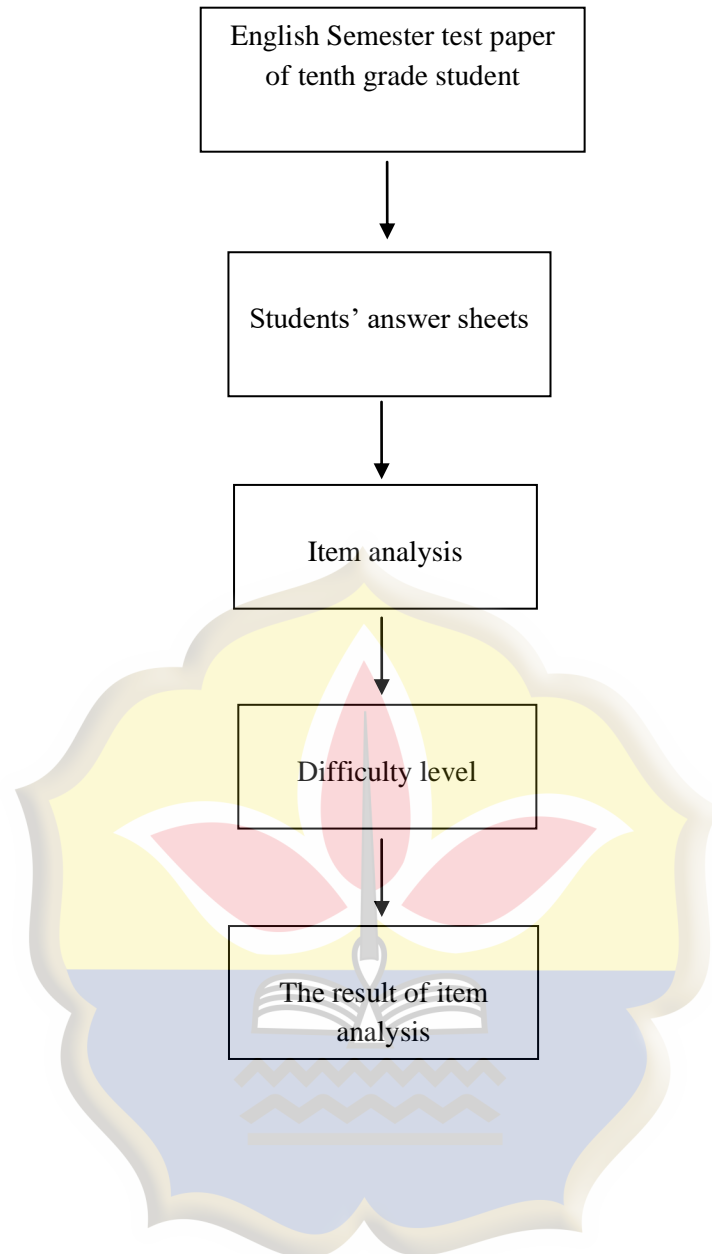
provided that necessary revision. From the results Boopathiraj and Chellamani concluded that developing and administering Multiple Choice Questions on the content knowledge of research methods in education helps teacher educators in molding future teachers.

Of the three previous studies above, there are similarities with the research will be conducted by the researcher, namely *an analysis on the quality of difficulty level items of english semester test at the tenth grade students of sman 3 jambi city in the academic year 2017/2018*. The researcher compares those studeis to her research that there is a similarity to the instrument that is in using the summative or semester test. Almost all of the studies are categorized as a descriptive analysis and considered as quantitative research because of using numerical data. The research can be categorized into descriptive analysis. This descriptive analysis is concerned with quantitative analysis. Quantitative is used in analyzing data of scores to detect the test item whether it is good or not. Then, in taking sample the researcher take for about 50 students' answer sheet as a sample.

2.4 Conceptual Framework

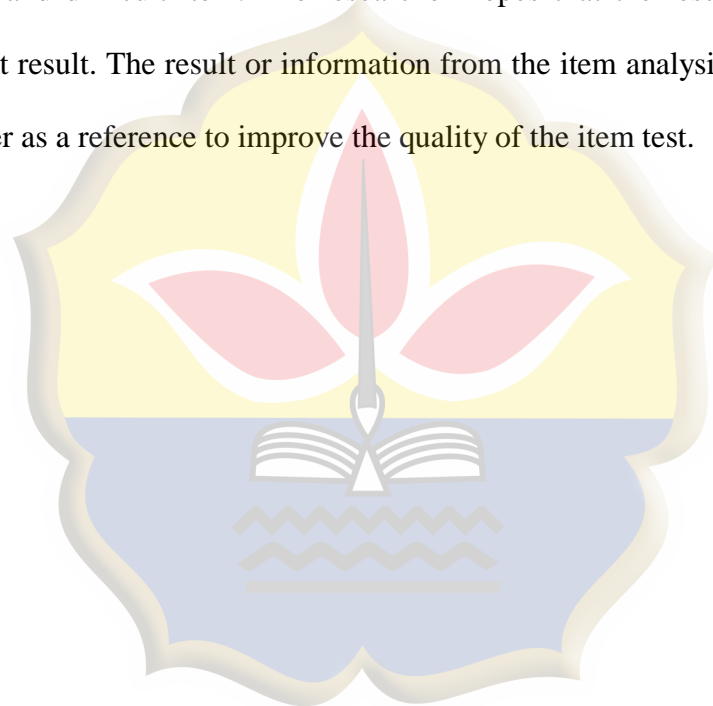
In this research the researcher just focus on the discuccion about “An Analysis on The Difficulty Level Items of English Semester Test at The Tenth Grade Students of Sman 3 Jambi City in The Academic Year 2016/2017”

Furthermore the researcher carried on the research based on the following framework:



First of all, the researcher do the preliminary during the researcher do *PPL* at SMAN 3 Jambi City for about three months. During *PPL* , the researcher observed the problem occurred at that school. The problem occurred that related to the examination test items. So, the researcher wanted to know the items that made by the teacher for the examination test from the difficulty level whether easy, moderate, or difficult.

The researcher used the documentation in conducting the research. The researcher took the data by asked the students' answer sheet, answer key and English semester test paper to the school. So, the documentations enclosed to get the data. After the conducting the data is succeed, then the researcher do the analysis of the data. The documentations as the data had been counted by using Heaton's formula and had been categorized into easy, moderate and difficult based on the score of each item. Then categorized the items, which one are easy, moderate and difficult item. The researcher hopes that the result will show the significant result. The result or information from the item analysis can be used by the teacher as a reference to improve the quality of the item test.



CHAPTER III

RESEARCH METHODOLOGY

3.1 Research Design

The shape of this research is descriptive quantitative research. Descriptive research is a research that conducted to determine the state and the development of a certain thing or phenomenon and describe something in detail and what it is (Singarimbun and Efendi, 1998:4). This research is descriptive because it is only to describe concrete about the quality of the item about the final test of English subject based on difficulty level of each item.

The meaning of quantitative research is a research that the process is required to use many numbers starting from the data collection process, data analysis, until the exposure of research results (Arikunto 2006:12) however, in addition beside the data that formed of the numbers in the research, there is qualitative as an explanatory.

This research is included in mix method research those are qualitative and quantitative research, in quantitative research because it takes the data research in the form of documents about the students' answers and questions are need to be counted to know the result of the data. Then, an analysis process of item of multiple choice test is based on difficulty level analysis need the accurate counting based on the formula which already exists. In qualitative because the results also are presented in the table and figures with explanatory words.

3.2 Population

In Sugiyono(2012:80), population is generalisation area which is consist of : object/subject which has current quality and characteristic that is stated by the

researcher to be studied and taken the conclusion from it. According to Arikunto (1998:115), population is whole research subject . from both experts above, the researcher can conclude that population is whole reseacrh subject that the quality and characteristic can be studied. The population in this research is in 10th grade in science classes. There are 7 classes.

Table 3.1 Population of The Research

Class	Number of Students
X MIA 1	40
X MIA 2	40
X MIA 3	42
X MIA 4	40
X MIA 5	43
X MIA 6	42
X MIA 7	30
Total Population	277

Source : Administration Office of SMA N 3 Jambi City Academic Year 2017/2017

3.3 Sample

According to Sugiyono(2014:118) “sample is part or the number and characteristic possessed by the population. The researcher takes the sample from the population is by using random sampling. Santoso said that (2007:49), random sampling is every unit or individual population who has the same probability to be a sample. The researcher uses simple random sampling. The researcher uses lottery to take the sample. Ii is choosen X MIA 2 as the sample.

3.4 Technique of Collecting Data

The instrument of the research to collect data is by using documentation. Documentation, derived from the word document, which means good written. In

carrying out the method of documentation, written objects such as books, magazines, documents, regulations, minutes meetings, diaries, and so forth (Arikunto 2006:158).

Documentation technique used in this research is for get data in the form of documents from the school. The data that formed of: a list of students' name, English semester test paper, students' answer sheets, and the answer key.

3.5 Technique of Data Analysis

Data analysis is the process of searching and compiling a systematic data obtained from interviews, field notes, and documentation, with organizing the data into categories, describe into the units, synthesizing, organize into a pattern, choose what is important and which will be studied, and make conclusions so it will be easily understood by yourself and others (Sugiyono, 2012). Data analysis is conducted after all the required data have been collected.

In this research, in analyzing the collected data, the researcher took the following steps:

- *Firstly*, the researcher took the English semester test paper and students' answer sheets from SMAN 3 Jambi city.
- *Secondly*, counted how many students answer the correct item based on the answer key of each item.
- *Thirdly*, counted the difficulty level by the formula.
- *fourthly*, categorized and calculated the result of the analysis of the English semester test for the tenth grade students at SMAN 3 Jambi city.
- *Lastly*, used *SPSS for Windows Version 16* to find out the validity of the items.

To calculate the level of difficulty, it can use Arikunto's (2013:223) formula as follows:

$$p = \frac{B}{JS}$$

P = Index of difficulty

B = The total number of person who get the item correct

JS = The number of students' who took a test.

Level of difficulty for each item must be interpreted in the rank scale of the difficulty level which are stated by Arikunto (2013:225).

Table 3.2 classification of difficulty level by Arikunto

- Item with P 0,00- 0,30 is difficult item
- Item with P 0,30-0,70 is moderate item
- Item with P 0,71-1,00 is easy item

After calculating the difficulty of each item, the writer calculates the index of difficulty of all item by Arifins's formula;

$$P = \frac{\sum b}{N}$$

P : difficulty level of all items

b : difficulty level of each items

\sum : sigma (total)

N : total numbers of test items.

To know the criteria of difficulty level of each item and all items, the researcher uses the measurement level referred to Suharsimi Arikunto's book. It can be seen on the previous page.

CHAPTER IV
FINDING AND DISCUSSIONS

4.1 Findings

In this chapter, the researcher revealed the data based on the research conducted during the research period. The data shown below gave the information about the difficulty level of each item.

4.1.1 Documentation

The documentation is conducted on Saturday, July 29th, 2017 at 09.00 in evaluation room. The researcher took 40 students' answer sheets of English examination test of X MIA 2 at SMAN 3 Jambi City in academic year 2016/2017.

Table 4.1 Result of Item Classification

NUMBER OF ITEM	KEY	TOTAL STUDENTS WHO ANSWERED CORRECT	$FV = \frac{R}{N}$	ID	INTERPRETATION	VALIDITY	INTERPRETATION
1	C	39	$\frac{39}{40}$	0,98	Easy	0,635	Valid
2	D	32	$\frac{32}{40}$	0,80	Easy	-0,105	Invalid
3	A	38	$\frac{38}{40}$	0,95	Easy	0,453	Valid
4	A	39	$\frac{39}{40}$	0,98	Easy	0,006	Invalid
5	E	35	$\frac{35}{40}$	0,88	Easy	0,337	Valid
6	C	24	$\frac{24}{40}$	0,73	Easy	0,048	Invalid
7	C	39	$\frac{39}{40}$	0,98	Easy	0,118	Invalid
8	C	39	$\frac{39}{40}$	0,98	Easy	-0,068	Invalid
9	E	4	$\frac{4}{40}$	0,10	Difficult	-0,067	Invalid
10	C	40	$\frac{40}{40}$	1,00	Easy	0,000	Invalid

11	E	33	33	0,83	Easy	-0.92	Invalid
			40				
12	D	37	37	0,93	Easy	0,415	Valid
			40				
13	E	38	38	0,95	Easy	0,549	Valid
			40				
14	D	37	37	0,93	Easy	0,336	Valid
			40				
15	A	38	38	0,95	Easy	0,549	Valid
			40				
16	B	40	40	1,00	Easy	0,000	Invalid
			40				
17	B	35	35	0,88	Easy	0,401	Valid
			40				
18	D	37	37	0,93	Easy	0,415	Valid
			40				
19	C	9	9	0,23	Difficult	0,004	Invalid
			40				
20	A	31	31	0,78	Easy	0,226	Invalid
			40				
21	A	10	10	0,25	Difficult	0,109	Invalid
			40				
22	D	40	40	1,00	Easy	0,000	Invalid
			40				
23	E	33	33	0,83	Easy	0,087	Invalid
			40				
24	A	23	23	0,58	Moderate	0,363	Valid
			40				
25	D	36	36	0,90	Easy	-0,067	Invalid
			40				
26	C	40	40	1,00	Easy	0,000	Invalid
			40				
27	A	40	40	1,00	Easy	0,000	Invalid
			40				
28	A	40	40	1,00	Easy	0,000	Invalid
			40				
29	C	40	40	1,00	Easy	0,000	Invalid
			40				
30	D	38	38	0,95	Easy	0,000	Valid
			40				
31	B	40	40	1,00	Easy	0,000	Invalid
			40				
32	A	40	40	1,00	Easy	0,000	Invalid
			40				
33	C	37	37	0,93	Easy	0,376	Valid
			40				

34	C	35	35	0,88	Easy	0,119	Invalid
			40				
35	A	38	38	0,95	Easy	-0,007	Invalid
			40				
36	C	32	32	0,80	Easy	0,089	Invalid
			40				
37	B	34	34	0,85	Easy	0,087	Invalid
			40				
38	D	40	40	1,00	Easy	0,000	Invalid
			40				
39	B	40	40	1,00	Easy	0,000	Invalid
			40				
40	C	15	15	0,38	Moderate	-0,166	Invalid
			40				

The result presented in table 4.2 revealed that there are 35 items are in easy level, they are in range 0,71 up to 1,00. There are 2 items are in moderate level, they are in range 0,31 up to 0,70. The rest of the item are 3 items in difficult level, they are in range 0,00 up to 0.30. Then, there is 12 valid item, and 28 invalid items.

4.2 Discussion

The English semester test consists of 40 multiple-choice items with five option. the researchet used 40 students' answer sheets, the researcher made the tabel of students' correct answer of each item. From the table in appendix 4 can be seen that there is no students can answer all the items corretly. But there are some items can be answered by students are correct all. They are in number 10, 40, 22, 26, 27, 28, 29, 31, 32, 38, and 39. Then, the number of item which get the lowest responden is in number 9, which are only 4 responden who correct the item.

After get the total of correct answer of the responden in each item. The researcher do the calculating to get the index of difficulty of each item. The way of calculating the difficulty level is the total number of person who get the item

correct then divided with the total of the number of students' who took a test. Then, from the result can be interpreted based on classification whether the range less than 0,30 it is concluded in difficult level, in the range 0,30 to 0,70 as moderate level, and the range more than 0,70 it is included in easy level.

After that researcher do the analyze of validity of the items. To know whether the item is valid or not. To get the validity if the item, the researcher used *SPSS for Windows Version 16*. From the appendix 5 can be seen the result of validity of each item. The researcher can conclude that there are 12 valid items, then there are 28 items are not valid. The valid items are in number 1, 3, 5, 12, 13, 14, 15, 17, 18, 24, 30, and 33. Then the items are not valid are in number 2, 4, 6, 7, 8, 9, 10, 11, 16, 19, 20, 21, 22, 23, 25, 26, 27, 28, 29, 31, 32, 34, 35, 36, 37, 38, 39, and 40.

From the data above, here are the results of counting the items by using formula to find the level of difficulty. Number 1 is categorized in easy level. Because there are 39 students from 40 students who can answer correctly. The difficulty level of this item is 0,98. It means that the range 0,71 to 1,00 that belongs to easy item. The item is valid with the value 0,635. Number 2 is categorized in easy level. Because there are 32 students from 40 students who can answer correctly. The difficulty level of this item is 0,80. It means that the range 0,71 to 1,00 that belongs to easy item. The item is invalid with the value -0,105. Number 3 is categorized still in easy level. Because there are 38 students can answer correctly. The difficulty level of this item is 0,95. It means that the range 0,71 to 1,00 that belongs to easy item. The item is valid with the value 0,453. Number 4 is categorized in easy level. Because there are 39 students from 40

students can answer the item correctly. The difficulty level of this item is in range 0,98 that belongs to easy item. The item is invalid with the value 0,006. Number 5 is categorized in easy level. Because there are 35 students from 40 students who can answer correctly. The difficulty level of this item is 0,88 that belongs to easy item. The item is valid with the range 0,337.

Number 6 is also categorized in easy level. Because there are 24 students from 40 students can answer the item correctly. The difficulty level of this item is 0,73 that belongs to easy item. The item is invalid with the value 0,048. Number 7 is categorized in easy level. Because there are 39 students from 40 students can answer the item correctly. The difficulty level of this item is 0,98 that belongs to easy item. The item is invalid with the value 0,118. Number 8 is categorized in easy level. Because there are 39 students from 40 students can answer correctly. The difficulty level of this item is 0,98 that belongs to easy item. The item is invalid with the value -0,068. Number 9 is categorized in difficult level. Because there are only 4 students from 40 students who can answer correctly. The difficulty level of this item is 0,10 that belongs to difficult item. The item is invalid with the value -0,067.

Number 10 is categorized in easy level. Because there are all students from 22 students answer correctly. The difficulty level of of this item is 1,00 that belongs to easy item. The item is invalid with the value 0,000. Number 11 is categorized in easy level. Because there are 33 students from 40 students who can answer the item correctly. The difficulty level of this item is in the range 0,83 that belongs to easy item. The item is invalid with the value -0,92. Number 12 is categorized in easy level. Because there are 37 students from 40 students can

answer the item correctly. The difficulty level of this item is in the range 0,93 that belongs to easy item. The item is valid with the value 0,415. Number 13 is categorized in easy level. Because there are 38 students from 40 students who can answer correctly. The difficulty level of this item is in the range 0,95 that belongs to easy item. The item is valid with the range 0,549. Number 14 is categorized in easy level. Because there are 37 students from 40 students can answer correctly. The difficulty level of this item is in the range 0,93 that belongs to easy item. The item is valid with the value 0.336.

Number 15 is categorized in easy level. Because there are 38 students from 40 students can answer correctly. The difficulty level of this item is in the range 0,95 that belongs to easy item. The item is valid with the value 0,549. Number 16 is categorized in easy level. Because all students can answer the item correctly. The difficulty level of this item is in the range 1,00 that belongs to easy item. The item is invalid with the value 0,000. Number 17 is categorized in easy level. Because there are 35 students from 40 students can answer the item correctly. The difficulty level of the item is in the range 0,88 that belongs to easy item. The item is valid with the value 0,401. Number 18 is categorized in easy level. because there are 37 students from 40 students can answer the item correctly. The difficulty level of the item is in the range 0,93 that belong to easy item. The item is valid with the value 0, 415.

Number 19 is categorized in difficult level. Because there are only 9 students from 40 students who can answer the item correctly. The difficulty level of the item of this item is in the range 0,23 that belongs to difficult item. The item is invalid with the value 0,004. Number 20 is categorized in easy level. There are

31 students who can answer the item correctly from 40 students. The difficulty level of this item is in the range 0,78 that belongs to easy item. The item is invalid with the value 0,226. Number 21 is categorized in difficult level. Because there are only 10 students from 40 students who can answer correctly. The difficulty level of this item is in the range 0,25 that belongs to difficult item. The item is invalid with the value 0,109. Number 22 is categorized in easy level. Because all students from 40 students can answer the item correctly. The difficulty level of this item is in the range 1,00 that belongs to easy item. The item is invalid with the value 0,000.

Number 23 is categorized in easy level. Because there are 33 students from 40 students can answer correctly. The difficulty level of this item is in the range 0,83 that belongs to easy item. The item is invalid with the value 0,087. Number 24 is categorized in moderate level. Because there are only 23 students who can answer correctly from 40 students. The difficulty level of this item is in the range 0,58 that belongs to moderate item. The item is valid with the value 0,363. Number 25 is categorized in easy level easy level. Because there are 36 students from 36 students answer the item correctly. The difficulty level of this item is in the range 0,90 that belongs to easy item. The item is invalid with the value -0,067. Number 26 is categorized in easy level. Because all students from 40 students answer the item correctly. The difficulty level of the item is in the range 1,00 that belongs to easy item. The item is invalid with the value 0,000.

Number 27 is categorized in easy level. Because all students from 40 students can answer correctly. The difficulty level of the item is in the range 1,00 that belongs to easy item. The item is invalid with the value 0,000. Number 28 is

categorized in easy level. Because all students from 40 students answer the item correctly. The difficulty level of the item is in the range 1,00 that belongs to easy item. The item is invalid with the value 0,000. Number 29 is categorized in easy level. Because all students from 40 students can answer correctly. The difficulty level of the item is in the range 1,00 that belongs to easy item. The item is valid with the value 0,000. Number 30 is categorized in easy level. Because there are 38 from 40 students can answer the question correctly. The difficulty level of the item is in the range 0,95 that belongs to easy item. The item is valid with the value 0,405. Number 31 is categorized in easy level. Because all the students from 40 students can answer correctly. The difficulty level of this item is in range 1,00 that belongs to easy item. The item is invalid with the value 0,000.

Number 32 is categorized in easy level. Because all the students from 40 students can answer correctly. The difficulty level of this item is in range 1,00 that belongs to easy item. The item is invalid with the value 0,000. Number 33 is categorized in easy level. Because there are 37 students from 40 students can answer the item correctly. The difficulty level of this item is in range 0,93 that belongs to easy item. The item is valid with the value 0,376. Number 34 is categorized in easy level. Because there are 35 students from 40 students can answer the item correctly. The difficulty level of this item in range 0,88 that belongs to easy item. The item is invalid with the value 0,119. Number 35 is categorized in easy level. Because all there are 38 students from 40 students can answer the item correctly. The difficulty level of this item in range 1,00 that belongs to easy item. The item is invalid with the value -0,007.

Number 36 is categorized in easy level. Because there are 18 students from 22 students answer the item correctly. The difficulty level of the item is in range 0,95 that belongs to easy item. The item is invalid with the value 0,089. Number 37 is categorized in easy level. Because there are 34 students from 40 students can answer the item correctly. The difficulty level of the item is in range 0,85 that belongs to easy item. The item is invalid with the value 0,087. Number 38 is categorized in easy level. Because there are 40 students from 40 students can answer correctly. The difficulty level of this item in range 1,00 that belongs to easy item. The item is invalid with the value 0,000. Number 39 is categorized in easy level because all students from 40 students can answer the question correctly. The difficulty level of this item is in range 1,00 that belongs to easy item. The item is invalid with the value 0,000. In the last number is number 40 is categorized in moderatet level. Because there are only 15 students from 40 student can answer the item correctly. The difficulty level of the item is in range 0,38 that belongs to moderate item. The item is invalid with the value -0,166.

Table 4.2 the category of difficulty level of the English semester test items

NO	Range of Difficulty Level	Category	Frequency
1	0,00-0,30	Difficult	35
2	0,30-0,70	Moderate	2
3	0,70-1,00	Easy	3
Total			40

From the table and the expalantion above, we can see that there are 35 items belongs to the easy item. They are items number 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18 ,20, 22, 23, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39. Which are in range 0,71 to 1,00. Then, there are 2 items belongs to

the moderate item. They are number 24,40. Which are in range 0,31 to 0,70. The other, there are 3 items belongs to the difficult item. They are number 9, 19, 21. Which are in range 0,00 to 0,30.

The last step is to count the difficulty level of all items by using the Arifin's formula, with the way sum the index of difficulty of each item then divided with the total number of test items.

$$P = \frac{\sum b}{N}$$

$$P = \frac{34,09}{40} = 0,85$$

From the calculation, can be seen the result. The researcher get the result is 0,85. The result of each item will be in decimal. As noted earlier, the researcher can interpret the result of difficulty level of all items according to Arikunto's criteria.

Therefore, the writer can interpret that the difficulty level of the English semester test which tested at SMA N 3 Jambi City is in easy level seen from the difficulty level of all items. It can be said so because it has difficulty level of all items 0,85 that is in range between 0,71 up to 1,00 like shown in previous chapter.

CHAPTER V

CONCLUSION AND SUGGESTION

5.1 Conclusion

Based on the discussion in the previous chapter, the researcher would like to conclude that the difficulty level of English semester test items for the tenth grade student of SMA N 3 Jambi City are as follows:

1. There are 35 items regarded as easy items because they are at easy level, in range from 0,71 up to 1.00.
2. There are 2 items regarded as moderate items, because they are in moderate level, in range from 0,31 up to 0,70.
3. There are 3 items regarded as difficult items because they are at difficult level, in range 0,00 up to 0,30.

Overall, from this analysis it can be said that semester test of English students for the tenth grade students at SMA N 3 Jambi City has easy level of difficulty level.

5.2 Suggestion

The researcher would like to give some suggestions concerning the item analysis result:

1. For the other researcher, based on the limitation of the study, further studies might be needed to explore this issue with larger aspects to research. Also, the other researchers To know whether an item difficult, moderate, and easy the teacher should analyze the students' test result.
2. For the student, hopefully this research is useful to them, that they can study more before do the test. They should be careful when do the the test

because the good test is the test that has good discriminating power and the difficulty level. So, they should read well the question when do the test.

3. For the reader, hopefully this research can be useful for them as a benchmarks and evaluation material, this research can add more knowledge for them about analyzing the test



REFERENCES

Books :

- Airasian, Peter W. 2008. *Classroom Assessment; Concept and Application 5th Edition*. Me Graw Hill. New York.
- Arifin, Zainal. (2011). *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosdakarya
- Arikunto, Suharsimi. (1998). *Prosedur Penelitian*. Jakarta. Rineka Cipta
- Arikunto, Suharsimi. (2004). *Prosedur Penelitian: Suatu Pendekatan Praktek*. Jakarta : PT. Rineka Cipta.
- Arikunto, Suharsimi. (2013). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara
- Bachman, Benjamin s. 1971. *Handbook on Formative and Summative of Students' Learning*. Longman. London.
- Brown, Dean, James. 1996. *Testing in Language Programs*. Prentice Hall Regents. Upper Saddle River.
- Djemari Mardapi. (2008). *Teknik Penyusunan Instrumen Tes dan Nontes*. Yogyakarta: Mitra Cendikia
- Gronlund, Norman E. 1968. *Constructing Achievement Tests*. Prentice Hill. Englewood Cliffs.
- Harrison, Andrew, 1983. *A Language Testing Handbook*. Macmillan Press. London.
- Henning, Grant. 1987. *A Guide to Language Testing*. Newbury House Publishers. Cambridge.
- Hughes, Arthur, *Testing for Language Teachers*, Cambridge: Cambridge University Press. 1989.
- Lado, Robert. 1961. *Language Testing*. Wing Tai Cheung Printing Co Ltd. Hong Kong.
- Madsen, Harold S. 1983. *Techiques in Testing*. Oxford University Press. Oxford.
- Namara, Tim Mc. 2000. *Language Testing*. Oxford University Press. Oxford.
- New york. Hopkins, Kenneth D., *Educational and Psychological measurement and Evaluation (8th Edition)*, University of Colorado, Boulder: Allyn & Bacon, 1998.
- Nitko, Anthony J. 1983. *Educational Test and Measurement, an Introduction*. Harcourt Brace Jovanovich, Inc. New York.

Noll, Victor H. 1965. *Educational Test and Measurement 2nd Edition*. Houghton Mifflin Company, Boston.

Oller, John W. 1979. *Language Tests at School. A Pragmatic Approach*. Longman. London.

Rea-Dickson, Pauline and Kevin Germain. 1992. *Evaluation*. Oxford University Press. New York.

Rummel, J Francis, H.H. Remmers, NL. Gage. 1960. *A Practical Introduction to Measurement and Evaluation*. Harper and Brothers Publishers. New York.

Sudijono, Anas. (2011). *Pengantar Evaluasi Pendidikan*. Jakarta: PT Raja Grafindo Persada

Sugiyono. (2014). *Metode Penelitian Pendidikan: Pendekatan Kuantitatif, Kualitatif, dan R&D*. Bandung : Alfabeta

Stanley, Julian C. 1964. *Measurement On Today's School*. Prentice Hill Inc. Englewood Cliffs.

Thoha, M. Chobib. 2003. *Teknik Evaluasi Pendidikan*, PT. Raja Grafindo Persada. Jakarta.

Tinambunan, Wilmar, *Evaluation of Student Achievement*, Jakarta: Departemen Pendidikan dan Kebudayaan. 1988.

Vallete, Rebecca M. 1997. *Modern Language Testing*. Harcourt Brace Jovanovich Publisher.

Journals:

Boophatiraj, C and Chellamani, K. (2013). *Analysis of Test Items on Difficulty Level and Discrimination Index in The Test for Research in Education*. International Journal of Science & Interdisciplinary Research, (Online), Vol. 2(2). <http://indianresearchjournals.com/pdf/IJSSIR/2013/February/15.pdf>, accessed on July 27th, 2017.

Khan, Gul-Ar Navi, et al. (2015). *Using Item Analysis in Essay Types Questions Given in Summative Examination of Medical : Facility Value, Discriminating Index*. International Journal of Research in Medical Science, (Online), Vol. 3 (1), 178-182. <http://www.ejmanager.com/mnstemp/93/93-1416410335.pdf>, accessed on July 27th, 2017.

E-Books :

Heaton, J.B. 1977. *Writing English Language Tests*, Longman. London, [and en.bookfi.net/book/1147043](http://en.bookfi.net/book/1147043)

- Amelia, Rika. (2010). *An Analysis of The English Summative Test Items in Terms of Difficulty Level*. Unpublished Thesis : Universitas Islam Negeri. Jakarta. <http://repository.uinjkt.ac.id/dspace/bitstream/123456789/5888/1/RIKA%20AMELIA-FITK>
- Ciptaningrum, Dwi. (2014). *An Item Analysis of English Summative Test on Difficulty Level and Discriminating Power*. Unpublished thesis : Universitas Islam Negeri. Jakarta. <http://repository.uinjkt.ac.id/dspace/bitstream/123456789/25421/3/DWI%20CIPTA%20NINGRUM-FITK.pdf>.
- Djemari, Mardapi. 2007. *Teknik Penyusunan Instrumen tes dan Nontes*, Internet: <Http://bawahan.co.cc/pembaharuan-pendidikan-/4profesionalisme/30pilgan.pdf>
- Karzuni. (2011). *Analisi Butir Soal Ulangan Akhir Semester (UAS) Mata Pelajaran Bahasa Indonesia SMK Kelas X Semester Gasal Tahun Ajaran 2010/1011(Studi Kasus di SMK Muhammadiyah Ungaran)*. Unpublished Thesis : Universitas Negeri Semarang. <http://lib.unnes.ac.id/11149/>.

