FINAL PROJECT

**ITEM ANALYSIS OF ENGLISH NATIONAL
EXAMINATION FOR JUNIOR HIGH SCHOOL
(A Case of SMP N 1 Sempor in the Academic Year 2017/2018)**

Submitted in partial fulfilment of the requirements
for the degree of
Sarjana Pendidikan in English

by
Eva Septiana Nurbaeti
2201414011

**ENGLISH DEPARTMENT
FACULTY OF LANGUAGES AND ARTS
UNIVERSITAS NEGERI SEMARANG
2019**

APPROVAL

This final project entitled *Item Analysis of English National Examination for Junior High School (A Case of SMP N 1 Sempor in the Academic Year of 2017/2018)* has been approved by the board of the examiners and officially verified by the Dean of English Department of Language and Arts Faculty of Universitas Negeri Semarang on August 2019.

Board of Examiners

1. **Chairperson**
   Dr. Sri Rejeki Urip, M.Hum.
   NIP. 196202211989012001

2. **Secretary**
   Bambang Purwanto, S.S., M.Hum.
   NIP. 197807282008121001

3. **First Examiner**
   Puji Astuti, S.Pd., M.Pd., Ph.D.
   NIP. 197806252008122001
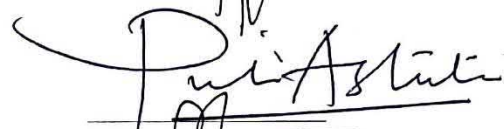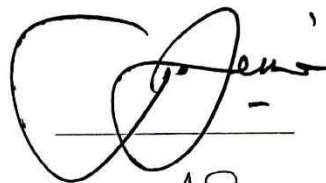
4. **Second Examiner**
   Pasca Kalisa, S.Pd., M.A., M.Pd.
   NIP. 198909062014042001

5. **Third Examiner as Advisor**
   Novia Trisanti, S.Pd., M.Pd.
   NIP. 197611062005012002

Approved by

Dean of Faculty of Language and Arts

Dr. Sri Rejeki Urip, M.Hum.
NIP. 196202211989012001

DECLARATION OF ORIGINALITY

I, Eva Septiana Nurbaeti, hereby declare that the final project entitled Item Analysis of English National Examination for Junior High School (A Case of SMP N 1 Sempor in the Academic Year of 2017/2018) is my own work and has not been submitted in any form for another degree or diploma at any university or another institute of tertiary education. Information derived from the published and unpublished work of others has been acknowledged in the text and a list of references is given in the bibliography.

Semarang, 30 Juni 2019

Eva Septiana Nurbaeti

iii

MOTTO AND DEDICATION

MOTTO:

"The best revenge is to make yourself better"

(Ali bin Abi Thalib)

DEDICATION:

This final project is dedicated to:

My beloved parents

My beloved brothers and sister

My spirit lifter

All of my best friends

ACKNOWLEDGEMENT

Semarang, 30 Juni 2019

Eva Septiana Nurbaeti

v

ABSTRACT

The purpose of this final project is to determine the quality of English National Examination for Junior High School in SMP N 1 Sempor in the Academic Year of 2017/2018 in terms of difficulty level, discrimination power, and distractor efficiency. The research design used by the writer in this study was a case study. The data used in this study were taken from the test papers and students' answer sheets. The writer took 32 students as the sample. The result of the research showed that: (1) the items that categorized as easy were 14 items (28%), moderate category were 31 items (32%), and difficult category were 5 items (10%). (2) the items which had poor discrimination power were 24 items (48%), satisfactory were 9 items (18%), good were 12 items (24%), and excellent were 5 items (10%). (3) the items that categorized to have not good distractors were 3 items (6%), fair distractors were 12 items (24%), good distractors were 17 items (34%), and very good distractors were 18 items (36%). The items that are included in the moderate category of difficulty level can be applied in the next test. The items which have poor discrimination power are needed to be revised, but the other categories can be used in the next test. In the analysis of distractor efficiency, the items that have very good, good, and fair distractors can be used in the next test. Meanwhile, the items with not good distractors should be revised. Based on the result, the writer suggests to the teachers and the test makers to prepare the test as good as possible.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF APENDICES

CHAPTER I

INTRODUCTION

In this first chapter, the writer would like to provide some subchapters. There are six subchapters, those are background of the study, reason for choosing the topic, research problem, significances of the study and outline of the study. Further discussion for the subchapters will be written below.

1.1 Background of the Study

Education is needed by everyone and it plays an important role in ensuring the continuity of the nation. Education is a way to improve the quality of human resources of a state. Human resources' quality can be developed by enhancing the quality of education in school. Education is a planned effort to develop the potential of students through a learning process. Education aims to develop students' potential to have the strength of spirit, personality, intelligence, and skills which can be implemented in society. With a good education, the students will have a better quality of human resources that will be the successor of the nation. The improvement of educational quality determined by various factors, such as teachers, students, environment, infrastructure, and learning process.

National examination is one part of Indonesian education process. English national examination is used to know the achievement of the English education in Indonesia. Since the national examination is a big event in Indonesian education, the quality of the items tested in the national examination must be good. To know the quality of the test, the teacher can conduct an item analysis activity. However,

based on the writer's dialogue with some teacher of junior high school, teacher usually used the previous national examination document as the try out to prepare the students in facing the next national examination. The teachers just distribute the document of the previous national examination without do any kind of analysis to know the quality of the previous national examination.

Analysis of the quality of a national examination is important to measure how far students' understanding of the materials that have been taught. Mostly, teachers just focus on the items without observing the quality of the test. Based on the case above, the teacher should do item analysis to know the quality of the test. Through the test item analysis, it will be obtained information about the good and the bad items. Item analysis can be done by calculating the aspect of validity, reliability, difficulty level, discrimination power, and distractor efficiency (Zainal Arifin, 2011). If a teacher knows the difficulty level, the discrimination power, and the efficiency of distractors, teacher can help students to improve their understanding and achievement in learning.

The limited ability of the teacher in some junior high school in doing item analysis mostly caused by there are too much steps that need to be done and the lack of time. The activity of the test item analysis takes a long time and considerable effort, thus that make some teachers did not do the item analysis. Eventhough, item analysis is an important activity to be conducted by the teacher to know the quality of the test items.

There are many kinds of test, one of them is teacher-made and standardized tests. A teacher-made test is a test made by the teacher in order to fill the needs of

assessment in a classroom. The effectiveness of this test depends on the competencies and abilities of the teacher in arranging a test. On the other hand, a standardized test is a test made by professional testing services to assist institutions in the selection, placement, and evaluation of students. An example of standardized test in Indonesia is the national examination or Ujian Nasional (UN).

.   National Examination or well known as UN is an evaluation system of the national standard for preliminary and secondary school, which done by the Ministry of National Education in Indonesia based on Undang-Undang Republik Indonesia no. 20 years 2003 stated that evaluation is held to control the quality of national education. Furthermore, the evaluation is carried out by the independent institution periodically, comprehensively, transparently and systematically to assess the achievement of national education standards and monitoring the process continuously.

National examination becomes an evaluation tool used to measure the achievement of education nationally, hence the test maker team must arrange the test items to have a good discrimination index, proportional difficulty level, and effective distractors which have high reliability. The test maker must have all of those characteristics so that the national examination can measure the achievement of education in Indonesia. Therefore, item analysis is needed to know how good the items to measure students' abilities.

According to the Ministry of Education and Culture (iNews, 2018), National Examination in 2018 are still in the form of multiple-choice questions. Multiple-choice question is commonly used in a test because it is effective, simple, and easy

to score. Knowing the proportion of the items tested in a national examination is important. Considering that the form of the test is multiple-choice questions, the distractor is one of the significant elements to define the quality of the test. Since the quality of the test is known, students are expected to improve their understanding of the material tested and prove it with their achievement with the good result of the national examination.

In this study, the writer analyzed the items of English National Examination for junior high school in the academic year of 2017/2018. The writer conducted the research in SMP N 1 Sempor. The writer chose this school because, in the academic year 2017/2018, this school still used the paper-based test for the national examination. The teacher said that the National Examination in the academic year of 2017/2018 is quite different, especially for English. The previous English National Examination consist of two skills, listening and reading, while in the academic year of 2017/2018, the English National Examination only consist of reading skill. The test consists of 50 items in the form of multiple-choice questions.

Since the test has not been analyzed by the teacher, this research will be very useful in helping the teachers to know the quality of the test. By considering the explanation above, the writer interested in conducting a research with the title: An Item Analysis of English National Examination for Junior High School in the Academic Year of 2017/2018.

1.2 Reasons for Choosing the Topic

The reasons why the writer was interested in choosing this topic above were based on several considerations:

1. The topic is relevant to the writer as part of the English Education Department.

2. The writer wants to know the quality of the items tested in the English National Examination.

3. By applying item analysis, we can indicate the difficulty level, index discrimination, and distractor efficiency of the English National Examination.

1.3 Research Questions

The writer formulates research problems as follow:

a. How is the difficulty level of reading skill in English National Examination for Junior High School?

b. How is the item discrimination power of reading skill in English National Examination for Junior High School?

c. How is the distractor efficiency of reading skill in English National Examination for Junior High School?

1.4 Objectives of the Study

The objectives of the study are:

a. To describe the difficulty level of reading skill in English National Examination for Junior High School.

b. To describe the item discrimination power of reading skill in English National Examination for Junior High School.

c. To describe the distractor efficiency of reading skill in English National Examination for Junior High School.

1.5 Significances of the Study

By conducting this research, the writer hopes that this research can provide some advantages as follows:

a. Theoretically, this study can give an explanation and a point a view of item analysis including difficulty level, discrimination power, and distractor efficiency.

b. Pedagogically, this study can help us in giving information about the quality of the English National Examination in terms of the difficulty level, discrimination, and distractor efficiency.

c. Practically, the result of this study will be useful for the teachers and test makers in improving their ability to construct the English National Examination better than before.

1.6 Outline of the Study

This study consists of five chapters. In the first chapter, the writer explains the background of the study, reason for choosing the topic, research questions, objectives of the study, significances of the study and outline of the study.

The second chapter presents the review of the previous study which provide some studies from other researcher that related with the topic of this study. Next

is review of the related literature. The writer will provide some theories that used in this study. The last sub-chapter is the research framework which explain how the study will be conducted.

The third chapter gives information about research methodology. This chapter consists of the research design, object of the study, and also the procedure of collecting the data and analyzing the data.

The next chapter deals with the explanation about the result of the research and the discussion. The writer explained the result of the analysis and compare the result of this study with the previous related study in discussion section.

The fifth chapter in this study shows the conclusion of the study and gives some suggestions. In this chapter, the writer will draw a conclusion based on the result of the analysis. The writer also gives suggestion based on the significances that stated in the chapter I.

CHAPTER II

REVIEW OF RELATED LITERATURE

In this chapter, the writer would like to provide some related researches and studies that have been done by other researchers. This part of the study was divided into three sections; review of previous studies which provide review of some researches done by other researchers, review of theoretical studies which provide some theories explained by experts that relevant to this study, and theoretical framework used by the writer to conduct this study.

2.1 Review of the Previous Studies

The first study was conducted by Anna and Freddano in 2011. This research aimed at investigating the effect of analysis of multiple-choice questions designed by the teacher on the quality of the test. The participants in this study are 108 students both male and female students from the 4th primary school which constituted the population of the study. The method used by the researcher was the classical approach. The result of the research showed that the level of difficulty was in the range between 0.26 and 0.75. The items were divided into four categories, easy, very easy, difficult and very difficult. Easy items consist of five items, very easy items consist of four items, difficult items consist of five items and the very difficult items consist of two items. The researcher suggested revision for question number 1, 3, 8, 11, 13, 15.

The second study was conducted by Boopathiraj in 2013. The study was aimed at analyzing test items of a researcher-made test in the subject of Research in Education for students of Master of Education. The participants were 200 students who taken randomly and used a test of multiple-choice of 60 items as a data collection instrument. The findings of this study showed that most items were in the acceptable range of level difficulty and discrimination level. However, some items with poor discrimination index were rejected.

The third study was conducted in 2014 by Dadan Rosana. The researcher published an article which aimed at getting a test item analysis program with Rasch model one parameter for testing the item difficulty level of a multiple-choice test and to determine the distribution of the item difficulty level of the test on grade XI students of Wonosari State Senior High School in the academic year 2013/2014. The result of the research showed that the difficulty level for the test with A code is about 7.5% very easy category, 12.5% easy, 25% moderate, 47.5% difficult, 2.5% very difficult, and 5% not good. Meanwhile, for code B is about 12.5% very easy, 32.5% easy, 45% moderate, 45% difficult, 2.5% very difficult, and 2.5% not good.

The next study was conducted by Anigbo in 2015. The study was conducted to investigate the psychometric properties of the National Examination Council (NECO)'s multiple-choice Economic test items for 2009-2011 academic sessions. The population of this research were nine thousand seven hundred and five students who sat for economic for three years under study. Survey is adopted as the research design in this study. The result of this study showed that no item achieved difficulty index under 0.3. On the other hand, 54 of 180 items (30%) have difficulty index

above 0.7. It showed that the difficulty index of the test was cheap. The discrimination index was below 0.3 which means weak discrimination. Besides, the performance of the distractors shows 88.7% effectiveness.

In the same year, Saputra conducted a study that aimed to analyze the quality of mid-term English test for seventh-graders which made by the state and private school certified teacher. The researcher applied item analysis to know the quality of the mid-term test in SMPN 1 Semarang and SMP Kesatrian 2 Semarang. The result shows that the English mid-term test in SMPN 1 Semarang categorized as an easy test with the average 0.775 of item facility index and 0.121 which means poor item discrimination power. On the other hand, the English mid-term test in SMP Kesatrian 2 Semarang categorized as an easy test with the average 0.799 of item facility index and 0.181 which means poor item discrimination power.

In the following year, Amrita Kumari conducted study deals with item analysis of diagnostic test in English language skills of the secondary school students. The analysis in this study involves the difficulty value and index of discrimination. The test consists of multiple-choice items to collect the data. Three hundred seventy students of the secondary school of Central Board of Secondary Education (C.B.S.E) were selected randomly for the sample. The result of this study showed that the items having difficulty level between 0.25 to 0.80 and above 0.25 for the discrimination power. From the total 86 items, 53 items were good, 1 item needs to be revised and 32 items were rejected. In terms of discrimination power, 40 items were categorized to be good items, 27 items need to be revised and 19 items were considered as poor and should be eliminated.

In the same year, Rusma Setiyana conducted a study about an analysis of English test at MAN Boarding School Meulaboh I. The data were taken from a checklist which analyzed using statistical procedures and the document that analyzed using Anates software version 4. The result showed that the index difficulty was above 70%, the discrimination index was 76% which means good, and the effectiveness of distractors was good.

In 2017, Toksoz and Ertunc conducted a study entitled Item Analysis of a Multiple-Choice Exam. This study aims to analyze the multiple-choice items aiming to test grammar, vocabulary, and reading comprehension administrated at a stated university student. 453 students' responses were analyzed in terms of item facility or difficulty index, item discrimination, and distractor efficiency. The result of the study reveals that most of the items are in the moderate level of item facility. Besides, the result showed that 28% of the items have low discrimination index. And last, the distractors were significantly ineffective and should be revised.

Still, in 2017, a study about analysis of the multiple-choice questions, including item difficulty, discrimination index, and distractor efficiency was conducted by Juliana and Maria. The aim of the research was to analyze 48 multiple-choice questions and 144 distractors and finding the relationship between the difficulty index and the discrimination index. The population in this research were 56 students and all the participant would be the sample of this research. This research showed that 29 items had an excellent difficulty index, 10 were good. All the distractors were well functioned. However, there was a significant negative correlation between the difficulty index and the discrimination index, the higher the

difficulty index, the ability to discriminate between the high and low achievers was decreased.

A year later, Danuwijaya (2018) studied item analysis of reading comprehension test for postgraduate students. The aim of this study is to analyze the reading comprehension test items. One hundred multiple-choice items were tested to 50 postgraduate students in a university. The result showed that the value of difficulty level is 0.47 which means the medium level of difficulty. On the other hand, 39% of the items have poor discrimination power, 23% items have medium discrimination power, 20% items have good discrimination power and only 18% items have excellent discrimination power. The test analyzed by the researcher consists of 100 items with four options; thus, the total number of distractors were 300. The researcher analyzed that 39 out of 300 of distractors were categorized as not-functioning.

The next study was a thesis done by Athiyah Salwa in 2012. In this study, the researcher wanted to present and compare the quality of two test packs involving the level of difficulty, discrimination power, and distractor distribution. The data were taken from students' grade V final tests and analyzed by using the descriptive comparative method. The findings of this study concluded that the whole test has good quality in validity, reliability, level difficulty, discrimination power, and distractor distribution. However, the test-pack 1 has a better quality that the test-pack 2, because there are some errors exist in test-pack 2.

Related with the previous study, Bernasela conducted research in 2014 which purposed to provide information about English test in Singkawang

Vocational School in the academic year 2012/2013. The researcher wanted to analyze the validity, reliability, level difficulty, discriminating power, and distractors of the items. The writer used the descriptive study as the research method. The data were taken from the test and the students' answer sheets. The whole analysis of this study concluded that 33 items are good and can be used for the next test, 6 test items which should be discarded and 11 test items that need to be revised.

In the same year, Dwi Ciptaningrum has done a study which aimed to measure the difficulty level and discriminating power of English summative test for the first grade of SMPN 3 Tangerang Selatan in the academic year 2013/2014. Dwi conducted qualitative research by using descriptive analysis to describe the difficulty level and discrimination power. The data were taken from 92 students and analyzed by using *Anates Program*. The result of this study showed that the value of the difficulty level on this test was 0.69 and categorized as a moderate level. Meanwhile, the value of discrimination power was 0.38 and categorized in good quality.

Still, in 2014, Teddy Fiktorius conducted study about the validation of the English National Examination of junior high school in Indonesia. The aim of the research was to provide feedback about the test quality from the local context that leads to the improvement of the English National Examination in Indonesia. The researcher used simple random sampling, which applied to obtain a packet of test items out of 42 packets. To know the quality of the national examination, the

researcher also done an item analysis, which consists of item difficulty, item discrimination, and the effectiveness of distractors.

The result of the study showed that the difficulty level of the test is neither too easy nor too difficult. Only 6.5% items which classified as good categorized, whereas the rest of the items are too easy and too difficult. The item discrimination had low discrimination value ranging from 0.71 to 1.00 because the examiners in both in the high and low groups can answer the items correctly. It means that the item did not differentiate students' abilities well. Finally, the distractor performances that the 50 items comprise 98 plausible distractors which chosen by at least 5% of total examinees, 52 implausible distractors which did not function well and even fail to attract any responses from the examinees, and 50 answer keys.

The next year, in 2015, Fatma Dwi Rusmiana conducted a study about test item analysis in the final test of accounting theory for SMK YPKK 1 Sleman. In this study, Fatma wanted to know the quality of the test in terms of level difficulty, discrimination index, and distractor efficiency. This study used a quantitative approach combined with the descriptive method to explain the result. The data were analyzed by using program Item and Test Program Analysis (ITEMAN) version MICROCAT 3.00. The result showed that 32 items (80%) indicated as the difficult category, 8 items (20%) medium category, and 0 (0%) items indicated as easy. The item discrimination power indicated as poor as much as 13 items (32.5%), enough 14 items (35%), good 8 items (20%), and the very good category were only 5 items (12.5%). Whereas for the distractor efficiency, very good category consists of 3

items (7.5%), 8 items had good enough (20%), enough 15 items (37.5%), bad 100 items (25%), and very bad consists of 4 items (10%).

In the following year, Ninuk Krismanti conducted a study about the analysis of listening test items of national examination try out in SMK 3 Semarang. The similarity of this study is equally analyzing test items consist of difficulty level and item discrimination power. The aim of this study was to reveal whether the test items have met the standard of a good test or not. The result of this study found that 14 items that being investigated, half of them belong to easy level and none of them belong to the difficult level. The difficulty level also regarding with the discrimination power, one of the 14 items tested has no discrimination power, and only one item has powerful discrimination power.

In 2016, Muspira Humaerah has done a study about item analysis of the English test for second-grade students of MAN 1 Ternate. As the topic stated, the purpose of this study was to know the validity, reliability, and difficulty level of the test. Humaerah applied a quantitative descriptive method to analyze the data obtained from the English test for social science class. Based on the whole analysis, it can be concluded that the test has good validity and the test was reliable. For the difficulty level, the test contains one difficult item, one too easy item, four medium items, and four easy items.

A study aimed of the study was to investigate the quality of multiple-choice items test created by teachers by Siregar was conducted in 2017. The test consists of 40 multiple-choice items and the classical theory was carried out. As the result, the difficulty level revealed that 17 items were easy, 4 items were moderate, and 1

item was difficult. The distractor analysis indicated that 21 items were good, and 1 item was very good of 22 valid items. Overall, 22 items out of 40 were categorized as good.

In the following year, 2018, Nurulanis has done a study aimed to know the quality of the English test in SMPN 1 Surakarta. Based on the analysis of the data, the test was categorized as an easy test. From 50 multiple-choice items, there are 17 very easy items, 15 easy items, 10 medium items, 5 difficult items, and 0 very difficult items. For the discrimination power, the researcher found that there are 0 excellent items, 1 good item, 16 satisfactory items, 31 poor items, and 2 very poor items.

The newest study in 2019 which conducted to investigate the difficulty and discrimination indices, and the distractor efficiency of the test for the freshman common course at Gondar CIE. 176 exam papers were analyzed in terms of difficulty index, discrimination power, and distractor efficiency. The result of this study showed that the test as a whole has a moderate difficulty level and good distractor efficiency. However, the exam was poor in discrimination power. Only one item has good discrimination power and one more item excellent in discrimination. 41.9% of the items were either too easy or too difficult.

2.2 Review of the Theoretical Studies

2.2.1 National Examination

According to H. A. R. Tilaar (2006), the national examination is governments' effort to evaluate education nationally by establishing standardize national education. The result of the national examination is used as a mapping tool for educational problems in order to develop national education policies. Whereas Syawal Gultom (2012) stated that national examination is a standard evaluation system for primary and secondary schools in Indonesia. According to Hari Setiadi, national examination is an assessment of learning outcomes by the government in order to assess the achievement of students' competencies nationally on certain knowledge and technology. Based on the opinions above, it can be concluded that national examination is an evaluation system held nationally, which aims as a mapping tool of educational problems in order to develop national education policies.

2.2.2 Item Analysis

Item analysis is a technique used to know the effectiveness of a test. Anas Sudijono (2011) said that analysis is the identification process of some items about to do to get feedback good repairment and revision about it. According to Thomson and Levitov in Dr. Mohammad Shakil, item analysis investigates the performances of items considered individually either related to some external criterion or in relation to the remaining items on tests. Besides, Nana Sudjana (2011) stated that item analysis is aimed to obtain a better quality of questions, the teacher will get a description of students' real achievement in learning. It can be concluded that item analysis is the identification activity of items to know students' ability by repair

quality of the test that will be arranged. Item analysis can be divided into three types, index of difficulty, index of discrimination, and distractor efficiency.

Here is the more explanation about the difficulty level, discrimination power, and distractor efficiency. Below will explain more about the part of the item analysis.

a. The Level of Difficulty

Susan (1997) declared in her journal that the level of difficulty is simply the percentage of students taking the test who can answer the question correctly. Heaton (1975) stated that the level of difficulty shows how easy or difficult the particular item proved in the test. This term means "how hard is this test?". To make a good test, it must be neither too difficult nor too easy. The level of difficulties is used to show students' highest level of achievement on an item.

According to Witherington in Anas Sudijono (2001) index of the level difficulty is 0.00 to 1.00. The computation of the level of difficulty (P) can be done by dividing the students who answer the question correctly (B) with the total number of students who participated in the test (JS). The lower P-value shows the higher the difficulty level.

Heaton (1975) stated that the index difficulty runs from 1.00-0.71 indicating easy category of question. Index difficulty which runs from 0.70-0.31 indicating the medium category of question. And index difficulty that runs at 0.30-0.00 indicating difficult category of question.

b. Item Discrimination Power

Heaton (1975) stated that item discrimination power indicates the extent to which the item discriminates between the students, separating their level. Item discrimination power can be said as the degree of students' different achievement levels on an item of the test.

According to Gronlund (1982) the computation of item discrimination index (D) can be done by subtracting the number of students in the lower group who get the right item (L) from the number of students in the upper group who get the right item (U) and divide by a half of the total number of students ($\frac{1}{2}$.T).

Whereas, Tinambunan said that discrimination power takes values from 0.00 to 1.00. the higher the D-value for an item, the better that item discriminated. Any item which has the D-value of 0.41 or above is considered as good at discriminating student differences. The D-values that reach 0.21 and 0.40 are usually considered to be satisfactory. The items with the lower value than thus range should be revised to make more effective discriminators.

In other words, we can conclude that item discrimination power can be used to separate students who already mastered the lesson and not. The higher the index of discrimination, the more capable the items to distinguish students' competence.

c. Distractor efficiency

English National Examination is a test that provides with multiple-choice test item. The multiple-choice question offering several possible answers in which there is only one answer is correct. Several wrong choices are known as distractors. The main purpose of distractors is to make the participants in the test interested to

select it. The smart students will avoid the distractor, but on the other hand, the less smart students will choose the distractor as their answer.

Susan (1997) stated that analyzing the distractor is useful in determining the relative usefulness of the decoys in each item. According to Anas Sudijono (2011), the distractors have done its function properly if it reaches 5% of the total participant of a test. The effectiveness of distractors is how many choices can distract the test participants who did not know the right answer. The more candidates choose the distractors, it means that the distractor is working well.
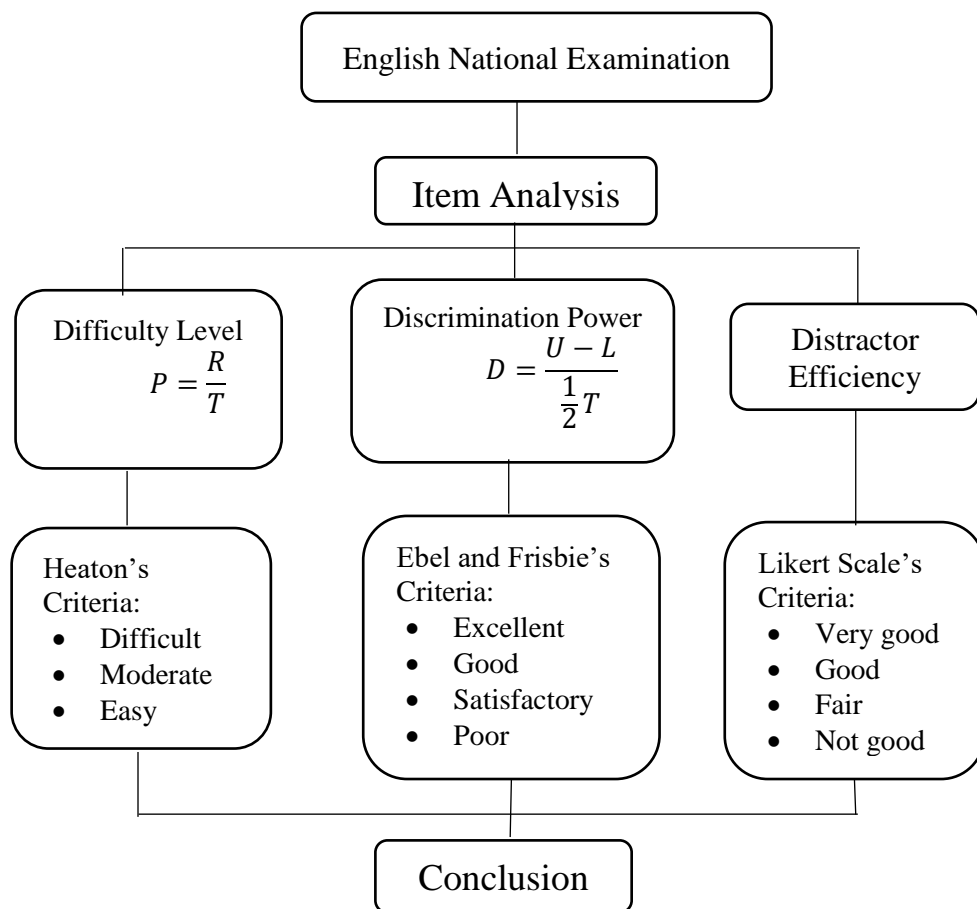
2.3 Theoretical Framework

Figure 2.1
Theoretical Framework

Item analysis is a systematic procedure that will provide specific information of a test. This study is intended to determine the quality of the items or questions in the English National Examination. The difficulty level is aimed to know whether the question is too difficult or not. Difficulty level divided into three categories, hard, moderate, and easy. The higher the difficulty level of an item, the fewer students who can answer the question correctly. While item discrimination power is used to determine the proficient and less proficient students. The item discrimination power divided into four categories, poor, satisfactory, good, and excellent. Then, the purpose of distractor efficiency is to attract students to select it. A distractor is declared to have a proper function if there are 5% of the students who choose it. The result of the item analysis will greatly help provide in-depth information regarding difficulty level, item discrimination power, and the distractor efficiency.

The item analysis is useful in analyzing the question in the English National Examination of SMPN 1 Sempor academic year 2017/2018 in order to know the quality of the item tested. The test result obtained from unqualified items, of course, cannot be a true reflection of students' achievement. The activity of item analysis will include the difficulty level, item discrimination power, and the distractor efficiency. This item analysis activity is aimed to provide information to the teacher about the quality of the item used. The teacher can find out the item quality and the result can be used to develop and revise the items which have less good or bad quality.

CHAPTER V

CONCLUSIONS AND SUGGESTIONS

In this chapter, the writer would like to provide conclusions, suggestions and also the limitation of the study. The conclusions support the finding and discussion in chapter IV which are the answer to the research questions. The suggestions support the significances of the study, which are dedicated to the parties mentioned in chapter I.

5.1 Conclusions

Based on the result of the analysis of the 50 test items of the English National Examination in SMP N 1 Sempor in the academic year of 2017/2018, the following conclusions could be drawn:

a. In the analysis of the difficulty level, the items that categorized to easy are 14 items (28%), moderate are 31 items (62%), and difficult are 5 items (10%). The items which are included in the moderate category can be used in the next test. Otherwise, easy and difficult items should be revised.

b. In the analysis of the discrimination power, the items categorized as poor are 24 items (48%), satisfactory are 9 items (18%), good are 12 items (24%), and excellent are 5 items (10%). In this case, the poor items need to be revised, so the poor items can differentiate the students in the upper and lower groups as well. The items with excellent, good, and satisfactory items can be applied for the next test.

c. In the analysis of the distractor efficiency, the items that categorized had not good distractors are 3 items (6%), fair distractors 12 items (24%), good distractors are 17 items (34%), and very good distractors are 18 items (36%). The items which have very good, good, and fair distractors still can be used in the next test. However, the items that have not good distractors should be revised.

In this study, the writer analyzed the item analysis of the test in terms of the difficulty level, discrimination power, and also the distractor efficiency as one of the types of item analysis. The weakness of this study is, in the process of analyzing the data, the writer did not know the validity and reliability of the test. According to Zainal Arifin (2011), to know the quality of the test, item analysis can be done by calculating the aspect of validity, reliability, difficulty level, discrimination power, and distractor efficiency. Thus, the writer should analyze the validity and also the reliability of the test, in order to know the whole quality of the test.

5.2 Suggestions

An item on a test should have good quality. To know the quality of the test, we can do item analysis including difficulty level, discrimination power, and distractor efficiency. Considering the conclusions above, the writer brings up some suggestions.

The first suggestion is for the teachers. It is important for the teacher and also teacher candidates to know the quality of the test. The teachers especially

English teacher could use this study as the consideration in order to do item analysis to know the quality of the test before given to the students.

The second suggestion is for the government. The government may be considering more the characteristic of a good national examination, so the next national examination will have better quality.

The last suggestion is for the future researcher. This study is limited to the object of the study. We know that the English National Examination has some packages, however, due to the lack of data sources the writer only analyzed one package. The future researchers are hoped to analyze two or more packages so they can compare the quality of each package. The writer also suggested to the future researcher to analyze the validity and reliability, not only the difficulty level, discrimination power, and distractor efficiency of the test.

5.3 Limitation of the Study

In this study, the writer analyzed the item analysis of the test in terms of the difficulty level, discrimination power, and also the distractor efficiency as one of the types of item analysis. The weakness of this study is, in the process of analyzing the data, the writer did not know the validity and reliability of the test. According to Zainal Arifin (2011), to know the quality of the test, item analysis can be done by calculating the aspect of validity, reliability, difficulty level, discrimination power, and distractor efficiency. Thus, the writer should analyze the validity and also the reliability of the test, in order to make the analysis more complete.

REFERENCES

Angelo, T. A. (1991). *Ten Easy Pieces: Assessing Higher Learning in four Dimension. New Direction in Teaching and Learning*, Summer 17-31.

Anigbo, L. C., & Ph, D. (2015). *Item Analysis of National Examination Council Senior School Certificate Examination Economics Objective Tests*, *3*(4), 23–30.

Arifin, Zaenal. (2009). *Evaluasi Pembelajaran.* Bandung: Rosda.

Arikunto, Suharismi. (2005). *Metode Penelitian Kualitatif.* Jakarta: Sagung Seto

Arikunto, Suharsimi. (2015). *Dasar-Dasar Evaluasi Pendidikan.* Jakarta: Bumi Aksara

Ary, D., Jacobs, L. C., & Soresen, C. (2010). *Introduction to Research in Education.* USA: Wadsworth

Bernasela. (2014). *An Analysis on English Summative Test Items*. Tanjung Pura University.

Boopathiraj, C. (2013). *Analysis of Test Items on Difficulty Level and Discrimination Index in the Test for Research in Education*, *2*(2), 189–193.

Brown, J. D. (2005). *Norm-Referenced Item Analysis (Item Facility and Item Discrimination)*. Shriken: JALT Testing & Evaluation SIG Newsletter 7(2). June 2003. (p. 16-19).

Ciptaningrum, D. (2014). *An Item Analysis of English Summative Test on DIfficulty Level and Discrimination Power*. Syarif Hidayatullah University, Jakarta.

Creswell, John W. (2009). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. Newbury Park: Sage Publication.

Danuwijaya, A. A. (2018). *Item Analysis of Reading Comprehension Test for Post-Graduate Students. 7*(1). https://doi.org/10.25134/erjee.v7i1.1493.Received

Ebel, R. L. & Frisbie, D. A. (1991). *Essentials of Educational Measurements.* 5th Edition. New Jersey: Pretince Hall, Inc.

Ferrando, Pere J. (2012). *Assessing the Discrimination Power of Item and Test Score in the Linear Factor-Analysis Model*. Rovira i Virgili University, Spain.

Fiktorious, Teddy. (2014). *Validation Study on National English Examination for Junior High School in Indonesia.* Thesis. Universitas Tanjung Pura, Pontianak.

Gay, L. R (1987). *Educational Research: Competencies for Analysis and Application, Third Edition.* Columbus Toronto, London and Melbourne: Merril Publishing Company.

Gronlund, N. E. (1982) *Measurement and Evaluation in Teaching 4th Edition.* Cliff, New York: Macmillan.

Gultom, Syawal. (2012) *Ujian Nasional Sebagai Wahana Evaluasi Pengembangan Pendidikan Karakter Bangsa.* Journal. p.5.

Haladyna, T. M. (2004). *Developing and Validating Multiple-Choice Test Items.* Mahwah, NJ, US: Lawrence Erlbaum.

Heaton, J. B. (1975) *Writing English Language Test.* London: Longman.

Humaerah, Musphira. (2016). *Item Analysis of English Summative Test for Second Grade of Students of MAN 1 Ternate Bulukumba.* Thesis. Universitas Alaudin Makasar.

Juliana & Mariana. (2017). *Analysis of Multiple Choice Questions : Item Difficulty , Discrimination Index.* (July). https://doi.org/10.5958/0974-9357.2017.00079.4

Krismanti, Ninuk. (2016). *The Analysis of Litening Test Items on National Examination Try-Out for 12th Graders of SMK 3 Semarang.* National Seminar Proceeding. STKIP Banjarmasin.

Kumari, A., & Bhattacharya, P. S. B. (2016). *Item Analysis of Diagnostic Test in English Language Skills of Secondary School Students, 6*(4), 1–5. https://doi.org/10.9790/7388-0604010105

Lutan, Rusli, dkk. (2000). *Tes dan Pengukuran Pendidikan.* Jakarta: Depdikbud.

Margono, Drs. (2003). *Metodologi Penelitian Pendidikan.* Jakarta: PT Asdi Mahasatya.

Marlow, C. (2001). *Research Method for Generalist Social Work.* Belmont, CA: Brooks/Cole Thompson Learning.

Matlock-Hetzel, Susan. (1997). *Basic Concept in Item and Test Analysis.* Pub date 23 Januari 1997. Accessed 27/07/2019

Nitko, J. Antony. (1983). *Educational Test Measurement: An Introduction.* New York: Marcourt Brace Jovanovich, Inc. p.6.

Nurulanis. (2018). *An Item Analysis of the English Summative Test on SMPN 1*

Surakarta. Undergraduate Thesis. Universitas Muhammadiah Surakarta.

Philips, Allen D. (1979). *Measurement and Evaluation in Physical Education.* Canada: John Whiley & Sons, Inc.

Rosana, Dadan. (2014). *Test Item Aalysis Program Develpoment With Rasch Model One Parameter for Testing the Item Difficulty Level of Multiple-Choice Test Using Bloodshed DEV C++ Applications.* Universitas Negeri Yogyakarta.

Rusmiana, F. Dewi. (2015). *The Test Item Analysis of 1st Semester Final Test of the Accounting Theory for Vocational Education: Case Study of SMK YPKK 1 Sleman Academic Year of 2014/2015.* Undergraduate Thesis. Yogyakarta State University.

Salwa, Athiyah. (2012). *The Validity. Reliability, Level of Difficulty, and Appropriateness of Curriculum of the English Test.* Thesis. Diponegoro University.

Saputra, R. W. (2015). Journal of English Language Teaching *The COmparison between the Second Mid-term English Test for the Seventh Graders Made by the State and Private School Certified English Teacher(The case of test items analysis of SMP N 1 Semarang and SMP Kesatrian 2 Semarang in the academic year of 2013 / 2014), 4*(1), 1–5.

Setiyana, Rusma. (2016). *Analysis of Sumative Test for English.* Undergraduate Thesis. Universitas Syiah Kuala. Banda Aceh.

Siregar, T. P., Surya, E., & Syahputra, E. (2017). *Quality Analysis of Multiple Choice Test and Classical Test at X Grade students of Senior High School.* International Journal of Advance Research And Innovative Ideas In Education (June).

Siri, A., & Freddano, M. (2018). *Procedia The use of item analysis for the improvement of objective examinations, 3*(February 2014). https://doi.org/10.1016/j.sbspro.2011.11.224

Sridandi. (2007). *Diktat Mata Kuliah Evaluasi Pembelajaran.* Yogyakarta: FIK UNY

Stufflebeam, D. L. (2003). *The CIPP Model for Evaluation.* Presented at the annual conference of the Oregn Program Evaluator Network (OPEN) 3 Oktober 2003. Diunduh di http://www.wmnich.edu/evalctr/cippmodel pada tanggal 26 Februari 2019.

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

Sudjana, Nana. (2011). *Penilaian Hasil Proses Belajar Mengajar.* Bandung: PT Remaja Rosdakarya.

Sugiyono. (2013). *Statistika Untuk Penelitian. Bandung*: Alfabeta.

Suryana. (2010). *Metode Penelitian: Model Praktis Penelitian Kuantitatif dan Kualitatif.* Bahann Ajar Perkuliahan UPI.

Tilaar, H. A. R. (2006). *Standarisasi Pendidikan Nasional: Suatu Tinjauan Kritis.* Jakarta: Rineka Cipta. (p 109-110).

Tinambunan, Wilmar. (1998). *Evaluation of Student Achievement.* Jakarta: Depdikbud.

Toksöz, S., & Ertunç, A. (2017). Item Analysis of a Multiple-Choice Exam, (c).

Yin, R. K. (2003). *Case Study Research: Design and Methods.* Thousand Oaks, Calif: Sage Publications.