



AN ANALYSIS OF TEACHER-MADE ENGLISH TRY-OUT TEST FOR NATIONAL EXAMINATION 2010-2011 FOR THE THIRD GRADERS OF MAN SIDOARJO

**Submitted in partial fulfillment of the requirement for the degree of
Sarjana Pendidikan Islam (S.Pd.I) In Teaching English**

PERPUSTAKAAN IAIN SUNAN AMPEL SURABAYA	
No. KLAS <i>K</i> T.2012 006 PB1	No. REG : T-2012/PB/1/06
ASAL BURU :	
TANGGAL :	

By

Abidatul Khoiro
NIM D05207050

ENGLISH EDUCATION DEPARTMENT
FACULTY OF TARBIYAH
STATE INSTITUTE FOR ISLAMIC STUDIES

SUNAN AMPEL SURABAYA

2012

GADJAHBELANG
8439407-5953789

APPROVAL SHEET

This thesis by Abidatul Khoiro entitled "*An Analysis of Teacher-Made English Try-Out Test for National Examination 2010-2011 for the Third Graders of MAN Sidoarjo*" has been approved by the thesis advisors for further approval by the Board of Examiners.

Surabaya, 30th January 2012

Advisor I

A hand-drawn diagram on a white background. It features two thin, horizontal black lines that run across the frame. A third, wavy black line starts from the left side, curves upwards and to the right, and then descends towards the right edge of the frame, where it meets another horizontal line.

Prof. H Arief Furqon , Ph.D
NIP. 150094096

Advisor II

6

Siti Asmiyah, M.Tesol
NIP.197704142006042003

APPROVAL SHEET

This thesis by Abidatul Khoiro entitled "*An Analysis of Teacher-Made English Try-Out Test for National Examination 2010-2011 for the Third Graders of MAN Sidoarjo*" has been examined on 9th of February 2012 and approved by the board of examiners.



The board of examiner
Chair,



Prof. H Arief Furqon , Ph.D
NIP: 150094096

Secretary,

Hilda Izzati Madjid, M.A
NIP.198602102011012012

Examiner I,

Akh. Muzakki, Grad Dip (SEA), M.Ag, M.Phil, Ph.D
NIP.197402091998031002

Examiner II,

Dr.Phil. Khoirun Ni'am
NIP.197007251996031004

ABSTRACT

Khoiro, Abidatul. 2012. *An Analysis of Teacher-Made English Try-Out test for National Examination 2010-2011 for the Third Graders of MAN Sidoarjo*. English Department, Faculty of Tarbiyah, State Institute for Islamic Studies Sunan Ampel Surabaya.

Advisor: Prof. H. Arief Furqon, Ph. D

Key words: Teacher-Made, Try-Out Test, National Examination, Content Validity, Index of Difficulty, Index of Discrimination.

Testing is very important in the teaching learning process. It is to measure the ability of the students. The English try-out test has an important role in measuring the students' ability and to face national examination. Thus, the teacher in MAN Sidoarjo constructed the test by their teacher selves. This study was conducted to know the quality of the teacher-made English try-out test of MAN Sidoarjo. It aims to know the content validity, index of difficulty and index of discrimination of the teacher-made English try-out test for national examination 2010-2011 for the third graders of MAN Sidoarjo.

The setting of the study was MAN Sidoarjo and the subject was the third class there is 40 students of Science class and 30 students of Social class. This study used descriptive method. The source of data in this study was documentation. The instrument was documents. The documents were applied in order to know the content validity, index of difficulty and index of discrimination of the teacher-made English try-out test.

The result shows that the content validity of the teacher-made English try-out test of MAN Sidoarjo has good content validity since 52% items test covered the indicators of Standard of Graduates Competencies. The test has acceptable index of difficulty because the Science class have 60% items which are adequate items and Social class have 68% items which are adequate items. And the index of discrimination of the test is different between both of class. The result of Science class shows that 44% items can be used. It means that the test is unacceptable for the Science class. And the Social class has acceptable index of discrimination since 60% items has satisfactory and good criteria.

From the finding and discussion, the suggestion for the test-maker and the teacher should analyze their own test, revise the items that too difficult and too easy and revise the poor and satisfactory item in order to discriminate between the students.

TABLE OF CONTENTS

COVER	i
APPROVAL PAGE	ii
MOTTO	iv
DEDICATION	v
ABSTRACT	vi
ACKNOWLEDGEMENT	vii
TABLE OF CONTENTS	viii
LIST OF TABLE	xi
APPENDICES	xii

CHAPTER I: INTRODUCTION

A. Background of the Study	1
B. Statement of the Problem	5
C. Objective of the Study	5
D. Significance of the Study	6
E. Scope and Limitation of the Study	6
F. Definition of Key Term	6

CHAPTER II: REVIEW OF RELATED LITERATURE

A. Review of the previous studies	8
B. Theoretical Background	10
1. Characteristic of a Good Test	10
a. Validity	10
1) Face validity	11
2) Content validity	12
3) Construct validity	13
4) Empirical validity	14
b. Reliability	14

c. Practicality	15
d. Authenticity	15
2. Item analysis	16
a. Index of difficulty	16
b. Index of discrimination	17

CHAPTER III: RESEARCH METHOD

A. Research Design	19
B. Research Setting	20
C. Subject of the Study	20
D. Data Collection Technique	20
E. Research Instrument	22
F. Data of the Study	22
G. Data Analysis Technique	23

CHAPTER IV: RESULT AND DISCUSSION

A. Description of the English Try-Out Test	25
B. The Result of Teacher-Made English Try-Out Test.....	26
1. XII Science Class	27
2. XII Social Class	29
C. Research Finding.....	30
1. Content Validity	30
2. Index of difficulty	34
3. Index of discrimination	40
D. Discussion	43
1. Content Validity	43
2. Index of difficulty	44
3. Index of discrimination	47

CHAPTER V: CONCLUSION AND SUGGESTION

A. Conclusion	50
B. Suggestion	51

BIBLIOGRAPHY 53

LIST OF TABLE

Table II.1	Example of analysis content validity.....	12
Table IV.1	Score list of XII Science Class	27
Table IV.2	Score list of XII Social Class	29
Table IV.3	Analysis content validity	31
Table IV.4	Analysis index of difficulty and index of discrimination of Science Class	35
Table IV.5	Analysis index of difficulty and index of discrimination of Social Class	37
Table IV.6	Criteria of index difficulty of XII Science Class	39
Table IV.7	Criteria of index difficulty of XII Social Class	40
Table IV.8	Criteria of index discrimination of XII Science Class	41
Table IV.9	Criteria of index discrimination of XII Social Class	42

CHAPTER I

INTRODUCTION

This study focuses on the analysis of the content validity, index of difficulty and index of discrimination of the teacher-made English try-out test for national examination. This chapter presents the background, the problems, the objectives, the significance of study, the scope and the limitation and the definition of key terms used in this study.

A. Background of the Study

Testing is an important part in every teaching and learning process. It is to measure the ability of the students and to know the progress of the students in learning process. According to Hughes, a test is an assessment to check the progress of the students and to measure what has been achieved both by groups and individuals.¹

A test is conducted for several purposes. It is useful for both the students who do the test and the teacher who administers the test. For the students, it can help them learn the language by requiring them to study hard and measure their ability. The test also shows students where they need to improve and the students will enhance their learning and are aware about objectives of the study if they are

¹ Arthur Hughes, *Testing for Language Teacher*, (Cambridge: Cambridge University Press, 1989), p.5

given a test.² It means that a test can motivate the students to study hard and show their ability in the subject of the study.

A test is also useful for the teacher. The teacher will know how far the students master the lesson and the difficulties they face. The test can help teachers to know the effectiveness of their teaching and to evaluate the teaching process.³ It means that a test makes it easier for the teachers to diagnose the students, so the test can help them to solve the students' difficulties and improve their teaching process.

In short, a test is very important to measure the students' achievements. Thus, students must pass a test when they want to finish their study. In Indonesia, the students in the third grade of high school have to pass the national examination or UN (Ujian Nasional).

The national examination is a test used to determine whether students are able to graduate from the school or not. The students in the highest grade of school must pass national examination; 6th grade of elementary school, 3rd grade of junior high school and 3rd grade of senior high school. Among subjects tested in national examination is English. Therefore, many schools prepare the students for the national examination by conducting tests or exercises called try-out. The try-out test it self has several purposes. First, it is to measure the readiness of the students

² Harold S Madsen, *Technique in Testing*, (New York: Oxford University Press, 1983), pp.4-5
³ Ibid, p.5

to face the national examination and to inform the teacher which competence that students need to improve. As an exercise to prepare students for the national examination, the try-out test must have high quality and validity relatively similar to the test in national examination.

The national examination and English try-out test has important role in measuring the students' ability. They must have the characteristic of a good test. Brown mentions that a test is a good test if it consists of five criteria of good test; validity, reliability, practicality, authenticity, and wash back.⁴ Validity is one of the important things to measure accuracy of the test. Among test validity is content validity. It is to measure the content of the test. The test has content validity if the test can represent sample of the content of the course. The purpose and material of the course is based on curriculum. In Indonesia, the material of national examination and try-out test is based on Standard of Graduates Competence. Because of it, English try-out test must be appropriate with Standard of Graduates Competence in order to have content validity.

Standard of Graduates Competence is guidance in assessment to decide on passing of the students in every level of school. Standard of Graduates Competence of national examination is developed by the government as the role of national education ministry number 46 year 2010 date December 31, 2010.

⁴ H. Douglas Bown, *Language Assessment: Principles and Classroom Practices*, (New York: Pearson Education, 2004), p.3.

Standard of Graduates Competence of English test concern to listening and reading skill.

There are two kinds of English try-out test based on the test maker. First, it is constructed by the government. The government conducts try-out test once or twice before the national examination, usually a month before national examination. This test is standardized test. Second, it is constructed by the teacher. One of the schools that construct English try-out test is MAN Sidoarjo.

MAN Sidoarjo is one of favorite schools in Sidoarjo because there is only one MAN or the state Islamic senior high school in Sidoarjo. Moreover, MAN Sidoarjo is one of many schools that has teacher-made try-out test. However, the teachers rarely analyze their own tests. In fact, some English try-out tests conducted by MAN Sidoarjo show students dynamic scores. Some students have good result in try-out but not in the national examination or vice versa or they have good score in one try-out and poor in others. So, it is still questioned whether the test is appropriate with Standard of Graduates Competence or whether the test is too difficult or too easy so that it is cannot discriminate the students' ability.

Because of the case above, the teacher-made English try-out tests for national examination 2010-2011 need to be analyzed. This research analyzes the content validity, item of difficulty and item of discrimination of the teacher-made English try-out test. So the title of this study is 'An Analysis of the Teacher-Made

English try-out Test for National Examination 2010-2011 for the Third Graders of MAN Sidoarjo'.

B. Statement of the Problem

Based on the background above, the research problems are formulated as follows.

1. What is the content validity of the teacher-made English try-out test in national examination 2010-2011 for the third graders of MAN Sidoarjo?
 2. What is the index of difficulty of the teacher-made English try-out test in national examination 2010-2011 for third graders of MAN Sidoarjo?
 3. What is the index of discrimination of the teacher-made English try-out test in national examination 2010-2011 for third graders of MAN Sidoarjo?

C. Objective of the Study

Based on the problems above, this study is intended to:

1. know the content validity of the teacher-made English try-out test in national examination 2010-2011 for the third graders of MAN Sidoarjo.
 2. know the index of difficulty of the teacher-made English try-out test in national examination 2010-2011 for the third graders of MAN Sidoarjo.
 3. know the index of discrimination of the teacher-made English try-out test in national examination 2010-2011 for the third graders of MAN Sidoarjo.

D. Significance of the Study

The result of this study is important especially for the following reasons. The result of analysis of content validity, index of difficulty and index of discrimination of the test can be useful for the teacher in order to construct the test better. This result is also useful for the students of English Education Department as reference to construct the test when they became a teacher later. Moreover, this study can inspire the future researcher to find the topic similar with this study or any further research.

E. Scope and Limitation of the Study

The scope and limitation of this study is English try-out test in national examination 2010-2011 for third graders of MAN Sidoarjo. The test is made by an English teacher in MAN Sidoarjo. The test consists of twenty five multiple choices questions.

F. Definition of Key Term

To avoid misinterpretation, these terms are defined as follows:

- ### 1. Content Validity

Content validity is defined as any attempt to show that the content of the test is a representative sample from the domain that is to be tested.⁵

⁵ Glenn Fulcher, *Language Testing and Assessment*, (New York: Routledge, 2007), p.6

In this study the content validity is defined as the validity of the item of the teacher-made English try-out test that represents the Standard of Graduates Competence.

2. Index of Difficulty :

According to Heaton, index of difficulty is to know how easy or difficult the particular items in the test.⁶

In this study the index of difficulty means the difficulty of teacher-made English try-out test item.

3. Index of Discrimination

Index of discrimination is the item of a test that can discriminate between the upper and the lower students.⁷ In this study, the index of discrimination is defined as the items of the teacher-made English try-out test that indicate the extent to which the item discriminate between the students.

⁶ J.B.Heaton, *Writing English Language Test*, (New York: Longman, 1975), p. 178

⁷ Burhan Nurgiyantoro, *Penilaian Dalam Pengajaran Bahasa Dan Sastra*,(Yogyakarta: BPFE Yogyakarta, 2001), p. 140

CHAPTER II

REVIEW OF RELATED LITERATURE

This chapter discusses some issues related to quality of test which became the focus of the research. The issues covered include characteristic of a good test and also items analysis.

A. Review of Previous Study

Some research with similar topic has analyzed the quality of the test. First is the research conducted by Anita Nur Rahma. This research analyzed the English summative test items for semester 1 2002-2003 of the first year students of SMU in Surabaya. The research analyzed objective test. It focused on the reliability, index of difficulty, index of discrimination and item distracter of the test. The result of the research concluded that the test have adequate reliability, poor index of discrimination and bad distracters. This test also categorized as easy test.¹

The second research is conducted by Sigit Eko Saputro who analyzed the teacher-made English final test items for semester 1 2005-2006 of the first year students of SMA 2 Magetan. The research analyzed both multiple choice and completion test. It focused on the content validity, reliability, index of difficulty, index of discrimination and item distracter of the test. The result of the research

¹ Anita Nur Rahma, ‘*an analyzed the English summative test items for semester I 2002-2003 of the first year students of SMU in Surabaya*’, Thesis S1, (Surabaya: perpustakaan UNESA, 2004).

concluded that the test have high content validity, adequate reliability, acceptable index of difficulty, did not have adequate index of discrimination and effective distracters.²

Iffah Mursyidah Mayangsari conducted the research in 2009. The research analyzed teacher-made formative English test in SMA 2 Muhammadiyah Sidoarjo. The research focused on the content validity, reliability, item difficulty and item discrimination. This research used descriptive research as design in the study. The result of the analysis concluded that the test has high content validity, adequate reliability, acceptable item difficulty and acceptable item discrimination.³

The other research entitled “An Analysis Content Validity and Item Analysis of Semester II English Final Test for Tenth Grade Students of SMAN 3 Sidoarjo” by Millatul Islamiyah in 2010 focused on multiple choice items. The samples of this study are X1, X2 and X3. The study analyzed content validity, index of difficulty, index of discrimination and item distracters. The design used in this study is a descriptive research. The result of the study concluded that the test had good content validity, acceptable index of difficulty, satisfactory index of discrimination and good distracters.⁴

² Sigit Eko Saputro, ‘*an analyzed the teacher-made test items in UAS semester 1 2005-2006 of the first year students of SMA 2 Magetan*’, Thesis S1, (Surabaya: perpustakaan UNESA, 2006).

³ Iffah Mursyidah Mayangsari, ‘*an analysis of UAS English test of second semester 2008/2009 by teacher-made English test in SMA 2Muhammadiyah Sidoarjo*’, Thesis S1, (Surabaya: perpustakaan IAIN Sunan Ampel, 2009).

The differences this study from the others is that this study analyzes English try-out test that made by the English teacher while the others analyzes English final test. Many researchers analyzed teacher made formative English test but never analyzed try-out that made by the English teacher. This study analyzes the teacher-made English try-out test and focus in the content validity, index of difficulty and index of discrimination.

B. Theoretical Background

This theoretical background discusses some theoretical background related to the quality of the test include the characteristic of good test and item analysis.

1. Characteristic of a Good Test

A test is an important instrument in teaching learning process to measure students' mastery on the materials. To know the affectivities of a test, it has criteria for testing a test. According to Brown, there are five criteria of a good test; validity, reliability, practicality, authenticity, and wash back.⁵

a. Validity

A test is classified to be valid if it measures accuracy what it is intended to measure. According to Heaton, validity of a test is the extent to which it measures what it is supposed to measure and nothing else.

⁴ Millatul Islamiyah, ‘*an analysis content validity and item analysis of semester II English final test for tenth grade students of SMAN 3 Sidoarjo*’, Thesis S1, (Surabaya: perpustakaan IAIN Sunan Ampel, 2010).

⁵ H. Douglas Bown, *Language Assessment: Principles and Classroom Practices*, (New York: Pearson Education, 2004), p.3.

Heaton mentions that there are four types of validity; face validity, content validity, construct validity and empirical validity.⁶

1) Face validity

Hughes states:

“A test is said to have face validity if it looks as if it measures what it is supposed to measure. Face validity is not scientific notion and is not seen as providing evidence for construct validity, yet it can be very important”.⁷

The test has face validity if the test looks right to other tester, teachers, moderator and test-takers. It means that face validity measured by subjective judgment.

Face validity will be high if the students or test-takers encounter some or the entire characteristic of good face validity, as follows:

- a) The test well-constructed and familiar format task,
 - b) the test is doable within the allotted time limit,
 - c) the items are clear,
 - d) the test have clear directions,
 - e) the test related to the course work,
 - f) a difficulty level that presents a reasonable challenge.⁸

⁶ J.B.Heaton, *Writing English Language Test*, (New York: Longman, 1975), p. 159

⁷ Arthur Hughes, *Testing for Language Teacher*, (Cambridge: Cambridge University Press, 1989), p.33

⁸ H. Douglas Brown, *Language Assessment*..... p. 27

2) Content validity

Content validity is defined as any attempt to show that the content of the test is a representative sample from the domain that is to be tested.⁹

Hughes states that a test is said to have content validity if its content constitutes a representative sample of the language skills, structures, etc. with which it is meant to be concerned.¹⁰

Test has content validity if the test is appropriate with the lesson that is taught in the class. Alderson states that to analyze content validity of the test, we have to compare the test with the test's specification such syllabus, curriculum or domain specification.¹¹

To analyze content validity, table below shows an example of domain specification in vocabulary test.¹²

Table II.1
Example of analysis content validity

Specification	Total Items	%
Verb	12	48
Noun	5	20

⁹ Glenn Fulcher, *Language Testing and Assessment*, (New York: Routledge, 2007). P.6

¹⁰ Arthur Hughes, *Testing for.....*, p.26

¹¹ J. Charles Alderson, et al, *Language Test Construction and Evaluation*, (New York: Cambridge University Press, 1995). P.173

¹² Prof. Dr. M. Soenardi Djiwandono, *Tes Bahasa Dalam Pengajaran*, (Bandung: ITB, 1996). P. 94

Pronoun	4	16
Phrase	3	12
Preposition	1	4
	25	100

As we see in the example above, proof of content validity is done by explanation, without using any counting and statistic. And to know if a test has high content validity or not, we can see from the percentage of the specification in the table. According to Milla's research, the test has high content validity if the agreement of the test is 50% or more.¹³

3) Construct validity

A test which has construct validity can measure certain specific characteristics concordant with a theory of language behavior. The type of validity supposes the existence of certain learning theories or constructs underlying the achievement of abilities and skills.¹⁴

Construct validity to show the score of the test is appropriate with the component being tested.

¹³ Millatul Islamiyah, *an analysis content validity and item analysis*p. 52
¹⁴ J.B.Heaton, *Writing English* p.161

4) Empirical validity

This validity is obtained as a result of comparing the results of the test with the results of some criterion measure such as:

- a) An existing test which is considered to be valid and given at the same time
 - b) The teacher's ratings or any other such form of independence evaluation given at the same time or later
 - c) The next performance of the tests must be measured by some valid test.¹⁵

b. Reliability

One of the necessary characteristic of good test is reliability. The test is said to be reliable if it is consistent in the measurements. It means that the students must have same mark if the test marked by two or more examiners.

Moreover, the reliability of the test is considering a number of factors that may contribute to the unreliability of the test. According to Heaton, the factors affecting the reliability are:

- 1) The extent of the material selected for testing.

¹⁵ Ibid, p.161

Reliability is concerned with the size of the test; it is not too long and not too short.

- ## 2) The administration of the test.¹⁶

The students or test-takers must have same condition and time limit.

c. Practicality

A test is said to have practicality if the test is not expensive, has appropriate time constraints, is easy to administered and easy to score.¹⁷ It means that the test does not consume more time and money and does not require many types of equipment. It also has efficient process of scoring.

Moreover, practicality is different with other qualities. While other qualities concern to the uses that are made of test scores, practicality concern to the way in implementing the test in a given situation.¹⁸

d. Authenticity

In a test, authenticity may be present in the following ways,

- 1) The language in the test is as natural as possible
 - 2) Items are contextualized rather than isolated
 - 3) Topics are meaningful for the learner

¹⁶ J.B.Heaton, *Writing English* p.162

¹⁷ H.Douglas Bown, *Language Assessment*....., p.28

¹⁸ Lyle F. Bachman, et all, *Language Testing In Practice*, (New York: Oxford University Press, 1996), p.35

- 4) Some thematic organization to items is provided, such as through a story line or episode
 - 5) Task represent or closely approximate, real world task.¹⁹

2. Items Analysis

The purpose of items analysis is to identify the test items whether it is good or not. To know the answer, all items should be identified from the index of difficulty and index of discrimination.

a. Index of difficulty

The good test items are not too easy and not too difficult. According to Heaton, index of difficulty is used to know how easy or difficult particular items in the test are. It is generally expressed as fraction or percentage of the students who answered the item correctly. To calculate the index of difficulty, Heaton uses the following formula.²⁰

$$FV = \frac{R}{N}$$

FV = Index of difficulty

R = the number of correct answer

N = the number of students taking the test

¹⁹ H. Douglas Brown, *Language Assessment*..... p. 28

²⁰ H. Douglas Brown, *Language Assessment*, J.B.Heaton, *Writing English* p. 178

Arikunto classifies the criteria of index of difficulty as follows.²¹

FV: 0,00 - 0,30 = difficult

FV: 0, 30 – 0, 70 = moderate

FV: 0, 70 – 1.00 = easy

It means that a good test to be given to the students is the test with the criterion index of difficulty between 0,30 – 0,70. Meanwhile, the index of difficulty which shows 0,00 – 0,30 and 0,70 – 1,00 is not good to be given to the students because the test is either too difficult or too easy for them.

b. Index of discrimination

Index of discrimination indicates the extent to which the items discriminate between the students. It is to discriminate the students who have high ability on the test and the students who have low ability on the test.²² Heaton's formula to calculate index of discrimination is:

$$D = \frac{\text{Correct U} - \text{Correct L}}{n}$$

D = Index of discrimination

Correct U = the number of upper group who answer correct

Correct L = the number of lower group who answer correct

n = the number of one group

²¹ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, (Jakarta: Bumi Aksara, 1996), p.214

²² J.B.Heaton, *Writing English* p.180

Arikunto classifies the criteria of index of discrimination as follows.²³

D: 0,00 – 0,20 = poor

D: 0,20 – 0,40 = satisfactory

D; 0,40 – 0,70 = good

D: 0,70 – 1,00 = excellent

The range index of discrimination according to Heaton as follows.

+1 = an item which discriminates perfectly

0 = an item which does not discriminate in any way at all

²⁴ -1 = an item which discriminates in entirely the wrong way.

²³ Suharsimi Arikunto, *Dasar-Dasar Evaluasi*..... p.223

²⁴ J.B.Heaton, *Writing English* p. 180

CHAPTER III

RESEARCH METHOD

This chapter presents the description of the research method that is used to analyze the content validity, index of difficulty and index of discrimination of the teacher-made English try-out test. The research method includes the research design, the research setting, the subject of the study, the data collection technique, the research instrument, the data of the study and the data analysis technique.

A. Research Design

This study used descriptive method. According to Danim, descriptive method is describe situation or phenomena or characteristic of individual or group accurately. Also, descriptive method is describing the condition of existence and classifying the information.¹ This method suits with this study because it describes the quality of the English try-out test with classified the data.

The study describes the quality of the teacher-made English try-out test in national examination 2010-2011 for the third graders of MAN Sidoarjo focusing the content validity, the index of difficulty and the index of discrimination.

¹ Sudarwan Danim, *Menjadi Peneliti Kualitatif*, (Bandung: Pustaka Setia, 2002), p. 41.

B. Research Setting

The study was conducted at MAN Sidoarjo which is located on Jl. Jenggolo (belakang stadion) number 02 Sidoarjo.

MAN Sidoarjo has chosen as the field of study because of the following reasons. MAN Sidoarjo is one of many schools that develop English try-out test by the teacher them selves, but the teachers rarely analyze their own test. In fact, some English try-out tests have been conducted by MAN Sidoarjo and the students have dynamic scores. So, it is still questioned whether the test has good quality or not.

C. Subject of The Study

The subject of the study is the third class of MAN Sidoarjo. There are two streams in MAN Sidoarjo, Science and Social streams. There are six classes of Science and four classes of Social streams. The number of the students at Science class is 180 students, and the number of the students at Social class is 120 students. Using random sampling technique, out of 300 students, 70 students were selected randomly from both streams; 40 are from Science class and 30 are from Social class.

D. Data Collection Technique

In this research, the source of data is documentation. According to Moleong, documentation is written data or films used to test and are interpreted. It is also as

a proof for experiment.² This study used the score of the students and the item test as the data; both are the documents. So the appropriate technique in this study is documentation.

Data of the research are collected through the following steps.

1. The data of the teacher-made English try-out test for national examination 2010-2011 and its answer key of are collected from the English teacher of MAN Sidoarjo who developed the test.
 2. The data of the students' answers sheet of the teacher-made English try-out test in national examination 2010-2011 are collected from the English teacher of MAN Sidoarjo who conducted the test.
 3. The data of the students' scores of the teacher-made English try-out test in national examination 2010-2011 are collected from the English teacher of MAN Sidoarjo who constructed the test.
 4. The data of Standard of Graduates Competence academic year 2010-2011 are collected from curriculum of KTSP that downloaded from www.ujiannasional.org.³

² DR.Lexy J.Moleong,M.A, *Metodologi Penelitian Kualitatif*, (Bandung: PT Remaja Rosdakarya,1989), p.161.

³ www.ujiannasional.org [accessed on 1st July 2011]

E. Research Instrument

This study used documents as instrument. According to Bogdan, there are 3 kinds of documents; personal documents, official documents and popular culture documents.⁴ This study used official document to answer the entire question of the study. The teacher-made English try-out test, the answer key in national examination 2010-2011 and Standard of Graduates Competence academic year 2010-2011 are used to answer the validity of the test. The students' answers sheet and the students' scores of the teacher-made English try-out test are used to answer the index of difficulty and index of discrimination of the teacher-made English try-out test.

F. Data of The Study

The data of this study are as follows:

1. The teacher-made English try-out test and the answer key in national examination 2010-2011 for third graders of MAN Sidoarjo.
 2. The students' answers sheet of the teacher-made English try-out test in national examination 2010-2011 for third graders of MAN Sidoarjo.
 3. The students' scores of the teacher-made English try-out test in national examination 2010-2011 for third graders of MAN Sidoarjo.
 4. Standard of graduates competence academic year 2010-2011

⁴ Robert Bogdan, *Qualitative Research for Education: An Introduction to Theories and Methods*, (New York: Pearson Education, 2006), p. 133.

G. Data Analysis Technique

There are some kinds of the data which must be analyzed. The data analyzed in this research are the content validity, the index of difficulty and index of discrimination of the try-out test. All the data collected are analyzed through the following.

1. Content validity

The content validity of the teacher-made English try-out test for national examination 2010-2011 for third graders of MAN Sidoarjo is analyzed by using the tables of specification based on Standard of Graduated Competencies. The steps are the following:

- a. Make the table of specification
 - b. Match the items in the test into the Standard of Graduated Competencies.
 - c. Calculate the percentage of the item
 - d. Conclude the result.

2. Index of difficulty

Before analyzing the index of difficulty, the students' score of teacher-made English try-out test is classified first.

- a. The steps to classify the students' score are:
 - 1) Arranged the students' score according to Science or Social class.
 - 2) Then, arranged the students' score from high to low
 - 3) Then, classified 50 % high score as upper group and 50 % low score as lower group became upper and lower group in each class.

- b. The steps to analyze the index of difficulty are:

 - 1) Make table analysis to ease in analyzing index of difficulty
 - 2) Then, compute the index of difficulty using the formula:

$$FV = \frac{R}{N}$$

FV = index of difficulty

R = number of students whose correct answer

N = number of students

- 3) Match the result of index of difficulty with the criteria.

3. Index of discrimination

And the last is analyzing the index of discrimination. The steps are:

- a) Make table analysis to ease in analyzing index of discrimination.
 - b) Then, compute the index of discrimination using the formula:

$$D = \frac{\text{Correct U} - \text{Correct L}}{n}$$

D = index of discrimination

Correct U = the number of students in upper group who answer the item correctly

Correct L = the number of students in lower group who answer the item correctly

n = number of candidate of one group

- c) Match the result of index of discrimination with the criteria.

CHAPTER IV

RESULT AND DISCUSSION

This chapter discusses the result of the research. Before the analysis of the content validity, index of difficulty and index discrimination of the test presented, this chapter describes the English try-out test first. Then, the result of the test, data analysis and the discussion of the teacher-made English try-out test in national examination 2010-2011 for the third graders of MAN Sidoarjo are presented in the last part of the chapter.

A. Description of the English Try-Out Test

Try-out test is a test to prepare students to face the national examination. MAN Sidoarjo is one of several schools that conduct try-out test to prepare students to face the national examination. MAN Sidoarjo conducted 7 times English try-out test for national examination 2010-2011. Five (5) tests are constructed by the teacher and 2 tests are constructed by the local government.

The students who did the try-out test are the third grade. There are two streams in MAN Sidoarjo, Science and Social. There are six classes of Science and four classes of Social. The number of the students at Science class is 180 students, and the number of the students at Social class is 120 students. So the number of students who did this tests approximately 300 students.

This study used the first Teacher-Made English Try-Out Test as an object of the study. The test and the other data were collected from the English teacher of MAN Sidoarjo on May 25th, 2011. This test was conducted on September, 22th 2010. It was conducted only an hour, at 2 till 3 pm. This test was constructed by the English teachers of MAN Sidoarjo. They used English try-out test from government and national examination test from last years as references. The test is objective test that consists of 25 multiple choice questions. The content of the test is reading skill.

B. The Result of the Teacher-Made English Try-Out Test

The subject of the study is the third graders of MAN Sidoarjo. There are two streams in MAN Sidoarjo, Science and Social streams. There are six classes of Science and four classes of Social streams. The number of the students at Science class is 180 students, and the number of the students at Social class is 120 students. Using random sampling technique, out of 300 students, 70 students were selected randomly from both streams; 40 are from Science class and 30 are from Social class.

After the students' scores were collected, the next step is arranging the students score according to Science or Social class. This makes the analyses of the test item easier.

1. XII Science Class

The next step after arranging the students score according to Science or Social class is arranging the students' score from high to low in each class. There are 40 students in XII Science class. To classify the upper and the lower students, the scores are divided half of the number of the students each group. Arikunto states that the small group of test-taker divided into 50% upper and 50% lower group, and for the big group it is divided into 27% upper and 27% lower group.¹ In XII Science, there are 20 students of upper group and 20 students of lower group as shown in the following table.

Table IV.1
Score List of XII Science Class

Upper Group		Lower Group	
No	Students' Score	No	Students' Score
1	84	1	60
2	76	2	60
3	76	3	60
4	76	4	60
5	72	5	60
6	72	6	60

¹ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, (Jakarta: Bumi Aksara, 1996), pp.216-217

7	72	7	60
8	72	8	60
9	68	9	56
10	68	10	56
11	68	11	56
12	68	12	52
13	68	13	52
14	64	14	48
15	64	15	44
16	64	16	44
17	64	17	40
18	64	18	36
19	64	19	32
20	60	20	28

Table above is used to make it easy to classify the upper and the lower students of Science class. The first column is the student's number who got high score, the second column is the score of the students who is classified in upper group, the third column is the student's number who got low score, and the last is the score of the students who is classified in lower group.

2. XII Social Class

In XII Social class there are 30 students. The step to divide students' score in upper and lower group is the same with Science class, 50% upper group and 50% lower group. So, in XII Social class, there are 15 students of upper group and 15 students of lower group as shown the following table.

Table IV.2
Score List of XII Social Class

Upper Group		Lower Group	
No	Students' Score	No	Students' Score
1	84	1	52
2	84	2	44
3	84	3	44
4	80	4	44
5	80	5	44
6	76	6	44
7	76	7	44
8	76	8	40
9	76	9	40
10	72	10	40
11	68	11	36

12	64	12	32
13	64	13	32
14	64	14	28
15	52	15	24

Arranged and classified students' score is used to analyze the index of difficulty and index of discrimination of the teacher-made English try-out test in national examination 2010-2011 for the third graders of MAN Sidoarjo.

C. Research Finding

After classifying the students to the upper to the low group, the next step is analyzing the content validity, index of difficulty and index of discrimination of the teacher-made English try-out test.

1. Content validity

The analysis of content validity of the teacher-made English test try-out for the third grades of MAN Sidoarjo uses Standard of Graduates Competencies 2011. It includes Standard competencies and indicators. Then, the materials of the test is matched with the Standard of Graduates Competencies to know the Teacher-Made English Test Try-Out for National Examination for the Third Graders of MAN Sidoarjo whether it fulfills the

agreement of content validity. The analyzing content validity used table specification (see table IV.3).

The table analysis is used to ease the analysis of content validity of the test. The table consists of 6 columns. The first column is number, the second column is General classification of Standard of Graduates Competencies, the third is indicators, and column number four, five and six consist of items number, sum of items and then percentage.

Table IV.3
Analysis Content Validity

No	Standart Kompetensi Lulusan	Indikator	Number of Items	Σ	%
1.	Mendengarkan Memahami makna dalam wacana lisan interpersonal 1 dan transaksional, secara formal maupun informal dalam bentuk teks fungsional pendek, <i>recount, narrative, procedure,</i>	Diperdengarkan sebuah percakapan lisan interpersonal/transaksional, siswa dapat menentukan gambaran umum/ informasi tertentu/informasi rinci tersurat dari percakapan tersebut. Diperdengarkan sebuah percakapan lisan interpersonal/transaksional, siswa dapat meresponnya/ melengkapinya dengan ungkapan rasa tak percaya/ puas/tidak puas/menyatakan/meminta pendapat/berjanji/ungkapan harapan.			
		Diperdengarkan sebuah percakapan lisan interpersonal/transaksional, siswa dapat menentukan gambar yang sesuai dengan percakapan			

	<i>descriptive, news item, report, analytical exposition, hortatory exposition, explanation, discussion dan review dalam konteks kehidupan sehari-hari.</i>	tersebut. Diperdengarkan sebuah teks monolog, siswa dapat menentukan gambar yang sesuai dengan teks tersebut.			
		Diperdengarkan sebuah teks monolog, siswa dapat menentukan gambaran umum/informasi tertentu/informasi rinci tersurat dari teks tersebut.			
2.	Membaca Memahami makna dalam wacana tertulis secara formal maupun informal dalam bentuk teks fungsional pendek, <i>recount, narrative, procedure, descriptive, news item, report, analytical exposition, hortatory exposition, explanation, discussion dan review</i> dalam	Disajikan sebuah teks fungsional pendek berupa <i>message</i> , siswa dapat menentukan gambaran umum/ rujukan kata/makna kata dalam teks tersebut.			
		Disajikan sebuah teks fungsional pendek berupa <i>announcement</i> , siswa dapat menentukan gambaran umum/informasi tertentu/informasi rinci tersurat dari teks tersebut.			
		Disajikan sebuah teks fungsional pendek berbentuk <i>advertisement/brochure</i> , siswa dapat menentukan gambaran umum/informasi tertentu/ makna kata dari dalam teks tersebut.	1,2,3	3	12
		Disajikan sebuah teks berbentuk <i>news item</i> , siswa dapat menentukan gambaran umum/informasi tertentu,informasi rinci tersurat/makna kata/frasa.	10,11, 12	3	12
		Disajikan sebuah teks berbentuk <i>recount</i> , siswa dapat menentukan informasi tertentu/pikiran utama paragraf/informasi	19,20, 21,22, 23,24, 25	7	28

	<p>konteks kehidupan sehari-hari.</p>	<p>tersirat/makna kata/frasa dari dalam teks.</p> <p>Disajikan sebuah teks berbentuk <i>explanation</i>, siswa dapat menentukan gambaran umum/informasi tertentu/informasi rinci tersurat/tujuan komunikatif/makna kata/frasa dari teks tersebut.</p> <p>Disajikan sebuah teks berbentuk <i>analytical/hortatory exposition</i>, siswa dapat menentukan gambaran umum/informasi tertentu/pikiran utama paragraf/informasi rinci tersurat/tersirat/makna kata/frasa dari dalam teks.</p> <p>Disajikan sebuah teks berbentuk <i>discussion</i>, siswa dapat menentukan gambaran umum/informasi tertentu/informasi rinci tersurat/informasi tersirat/makna kata/frasa dari teks tersebut.</p> <p>Disajikan sebuah teks berbentuk <i>review</i>, siswa dapat menentukan gambaran umum/informasi tertentu/informasi rinci tersurat/makna kata/frasa dari dalam teks.</p>		
	Total			13

The table above shows that there are two general classification of Standard of Graduates Competencies, listening and reading. In listening skill, there are five indicators, and there are nine indicators for reading.

Based on the result of the analysis of content validity in the table IV.3, all items are reading skill. Out of 9 indicators of reading skill, there are only 3 indicators included in the test. The percentage of indicators is 52% out of 25 items test for reading skill. It can be explained as follows:

- a. there are 3 or 12% items focus on advertisement/brochure text,
 - b. there are 3 or 12% items focus on news items text,
 - c. there are 7 or 28% items focus on recount text,

And there are 12 or 48% out of 25 items test that did not cover the indicators of Standard of Graduates Competencies. These items suitable with General classification of Standard of Graduates Competencies but they did not cover the indicators. There are 3 items focusing on narrative text, 3 items focusing on descriptive text, and 3 items focusing on report text.

2. Index of difficulty

Index of difficulty is explanation about how easy or difficult the test items for the students. Good items have adequate index of difficulty. It is not too easy or too difficult².

² Burhan Nurgiyantoro, *Penilaian Dalam Pengajaran Bahasa Dan Sastra*, (Yogyakarta: BPFE Yogyakarta, 2001), p. 138

After classifying the students' score to upper and lower group, the next step is analyzing the index of difficulty. The analysis the index of difficulty

uses Heaton's formula: $FV = \frac{R}{N}$

Note : FV = Index of difficulty

R = the number of correct answer

N = the number of students taking the test

The table is used to ease the analysis of the index of difficulty and index of discrimination.

Table IV.4
Analysis of index difficulty and index of discrimination of Science Class

Items No	U	L	U + L	U - L	F.V	D
1	19	18	37	1	0.93	0.05
2	11	5	16	6	0.40	0.30
3	17	15	32	2	0.80	0.10
4	19	17	36	2	0.90	0.10
5	18	13	31	5	0.78	0.25
6	11	11	22	0	0.55	0.00
7	9	8	17	1	0.43	0.05
8	12	10	22	2	0.55	0.10

9	16	13	29	3	0.73	0.15
10	10	2	12	8	0.30	0.40
11	14	9	23	5	0.58	0.25
12	18	9	27	9	0.68	0.45
13	14	7	21	7	0.53	0.35
14	16	16	32	0	0.80	0.00
15	12	5	17	7	0.43	0.35
16	4	5	9	-1	0.23	-0.05
17	16	9	25	7	0.63	0.35
18	14	13	27	1	0.68	0.05
19	18	17	35	1	0.88	0.05
20	20	13	33	7	0.83	0.35
21	13	11	24	2	0.60	0.10
22	17	9	26	8	0.65	0.40
23	9	7	16	2	0.40	0.10
24	10	3	13	7	0.33	0.35
25	10	11	21	-1	0.53	-0.05

Table IV.5
**Analysis of index difficulty and index of discrimination of
Social Class**

Item No	U	L	U + L	U - L	F.V	D
1	14	15	29	-1	0.97	-0.07
2	11	4	15	7	0.50	0.47
3	10	7	17	3	0.57	0.20
4	15	14	29	1	0.97	0.07
5	9	7	16	2	0.53	0.13
6	10	2	12	8	0.40	0.53
7	6	3	9	3	0.30	0.20
8	12	4	16	8	0.53	0.53
9	11	7	18	4	0.60	0.27
10	1	0	1	1	0.03	0.07
11	13	4	17	9	0.57	0.60
12	14	5	19	9	0.63	0.60
13	12	2	14	10	0.47	0.67
14	15	12	27	3	0.90	0.20
15	9	2	11	7	0.37	0.47
16	3	3	6	0	0.20	0.00
17	14	7	21	7	0.70	0.47

18	15	5	20	10	0.67	0.67
19	15	10	25	5	0.83	0.33
20	14	7	21	7	0.70	0.47
21	9	8	17	1	0.57	0.07
22	14	6	20	8	0.67	0.53
23	7	1	8	6	0.27	0.40
24	11	3	14	8	0.47	0.53
25	11	9	20	2	0.67	0.13

The tables above are used to analyze index of difficulty and index of discrimination. The first column is the items number of the test. There are 25 items number. The second column or (U) is the number of students in upper group who answer the item correctly. The third column or (L) is the number of students in lower group who answer the item correctly. The fourth column or (U+L) is the number of students in upper and lower group who answer the item correctly. The fifth column or (U-L) is the number of students in upper minus the number of students in lower group, who answer the item correctly. The sixth column or (F.V) is index of difficulty, it is the number of students in upper and lower group who answer the item correctly (U+L) divided the number of students taking the test. In science class, (U+L) divided 40 students, meanwhile, in social class; (U+L) divided 30 students. And the last

or (D) is index of discrimination. It is the numbers of students in upper minus the number of students in lower group who answer the item correctly (U-L) divide the number of one group. In science class, (U-L) divided 20 students and in social class (U-L) divided 15 students.

After analyzing the index of difficulty, the next step is matching the result with the criteria of index of difficulty according to Arikunto. The analysis is organized in the following table.

Table IV.6
Criteria of index difficulty
XII Science Class

Index of Difficulty	Criteria	Item Number	Total of Item
0, 00 – 0, 30	Difficult	10, 16	2
0, 30 – 0, 70	Moderate	2, 6 ,7, 8, 11, 12, 13, 15, 17, 18, 21, 22, 23, 24, 25	15
0, 70 – 1.00	Easy	1, 3, 4, 5, 9, 14, 19, 20	8

The table above shows that there are 15 items has moderate level of difficulty. The rest 10 items are either too easy or too difficult. And the following table is the table criteria of Social class.

Table IV.7
Criteria of Index Difficulty
XII Social Class

Index of Difficulty	Criteria	Item Number	Total of Item
0, 00 – 0, 30	Difficult	7, 10, 16, 23	4
0, 30 – 0, 70	Moderate	2, 3, 5, 6, 8, 9, 11, 12, 13, 15, 17, 18, 20, 21, 22, 24, 25	17
0, 70 – 1.00	Easy	1, 4, 14, 19	4

The table above shows that 17 items are moderate, 4 items are difficult and 4 items are easy.

3. Index of discrimination

Index of discrimination is used to show differences of ability between students who have high ability and the students who have low ability.

This study uses Heaton's formula to analyze index of discrimination of the teacher-made English try-out test in national examination 2010-2011 for the third graders of MAN Sidoarjo as follows:

$$D = \frac{\text{Correct U} - \text{Correct L}}{n}$$

Note:

D = Index of discrimination

Correct U = the number of upper group who answer correct

Correct L = the number of lower group who answer correct

n = the number of one group

The steps to analyze index of discrimination are the same with analyzing index of difficulty. The scores of the students were arranged and classified into the upper and lower students. The same table of index of difficulty to analyze index of discrimination is used.

After analyzing the index of discrimination, a table of criteria is used to match index of discrimination with criteria according to Arikunto.

Table IV.8
Criteria of Index Discrimination
XII Science Class

Index of Discrimination	Criteria	Item Number	Total of Item
0,00 – 0,20	Poor	1, 3, 4, 6, 7, 8, 9, 14, 18, 19, 21, 23	12
0,20 – 0,40	Satisfactory	2, 5, 11, 13, 15, 17, 20, 24	8
0,40 – 0,70	Good	10, 12, 22	3

0,70 – 1,00	Excellent	-	-
-0	Wrong	16, 25	2

The table above shows that there are 11 items have either satisfactory or good level of discrimination while the other 14 items are either poor or wrong discriminators.

Table IV.9
Criteria of Index Discrimination
XII Social Class

Index of Discrimination	Criteria	Item Number	Total of Item
0,00 – 0,20	Poor	3, 4, 5, 7, 10, 14, 16, 21, 25	9
0,20 – 0,40	Satisfactory	9, 19	2
0,40 – 0,70	Good	2, 6, 8, 11, 12, 13, 15, 17, 18, 20, 22, 23, 24	13
0,70 – 1,00	Excellent	-	-
-0	Wrong	1	1

The table above shows that 60% (15 out of 25 items) items have satisfactory and good criteria as discriminators. And the 10 items are either poor or wrong discriminators.

D. Discussion

From the finding above, the discussion of the content validity, index of difficulty and index of discrimination of the teacher-made English try-out test explained as follows.

1. Content Validity

Based on the result above, it concluded that the teacher-made English test try-out for the third grades of MAN Sidoarjo has 52% items test covered the indicators of Standard of Graduates Competencies.

And it is 48% items did not cover the indicators of Standard of Graduates Competencies. These items suitable with General classification of Standard of Graduates Competencies but they did not cover the indicators.

The conclusion of the result is the content validity of the teacher-made English try-out test of MAN Sidoarjo has good content validity since 52% items test covered the indicators of Standard of Graduates Competencies.

Although the test has good content validity, but there are still 12 out of 14 indicators of Standard of Graduates Competencies that not cover the test. It is caused by several possibilities; first, the distribution of the items is not

equal. There are 7 items in one indicator and the rest only 3 items in each indicator. So, if the distribution of the item test equal, it can cover more than 3 indicators. Second, the number of the question is only 25 question, it is very different with the national examination that 50 questions. So, it caused several indicators not cover the test. Last, some of the items not suitable with the indicator, but it is suitable with the general classification of the standard graduates of competencies. The possibility is the teachers refer only to the last national examination and not refer to the standard graduates of competencies in that year. So, in spite of the test has content validity, the test should be revise in order to the students more prepare to face national examination.

2. Index of Difficulty

a. Index of difficulty XII Science Class

According to the table IV.6, the result reported that there are 8 out of 25 items are easy, because they have index of difficulty value between 0,73 – 0,93. The items are number 1, 3, 4, 5, 9, 14, 19, and 20. These items can not be used because it is too difficult for the students to answer it.

And the items number 2, 6, 7, 8, 11, 12, 13, 15, 17, 18, 21, 22, 23, 24, and 25 have index of difficulty value between 0,33 – 0,68. These

items are categorized good or moderate items based on Arikunto's classification. The good items could be used in the test without revised.

While, the rest are difficult items, they are number 10 and 16. They have index of difficulty value between 0,23 – 0,30. These items can not be used because they could desperate students to answer these items. Difficulty items may also spend much time for the students to focus.

From the result above, it can be concluded that most of items or 15 out of 25 items are good items. They have difficulty value around 0,33 – 0,68. Arikunto stated that items have index of difficulty value between 0,31 – 0,70 are good. It means that this test has good index of difficulty.

b. Index of difficulty XII Social Class

The finding above showed index of difficulty XII Social Class relatively is moderate because there are 17 out of 25 items which they have index of difficulty value is approximately 0,37 – 0,70. The items are 2, 3, 5, 6, 8, 9, 11, 12, 13, 15, 17, 18, 20, 21, 22, 24, and 25. These items are proper to be offered to the students. Arikunto state that the good item test is moderate item, which is not too easy and not too difficult.³

Items 7, 10, 16, 23 have index of difficulty value between 0, 23 – 0, 30. It means that these numbers are classified in difficult category. This

³ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, (Jakarta: Bumi Aksara, 1996), p.214

category tends to burden the students because of the unsuitability between ability and quality of the question. Djiwandono stated that items which are very difficult can caused test-taker can not answer; that is useless.⁴

And the rest items are easy classification. It is still found the items that are categorized as easy. The items are 1, 4, 14, 19 which have index difficulty value between 0,73 – 0,93. Most of students can answer all of questions which have easy category; it does not give information about the students which have high ability and the students which have low ability.

In conclusion, the questions of the test are mainly categorized as moderate category because the numbers accomplished 17 out of 25. It means that the questions are precise to be tested to the student. These items can be used for the next test and also used to references for the test maker to construct other test.

Finally, it concluded that the teacher-made English try-out test have good index of difficulty, because both Science and Social class have 15 items or 60 % and 17 items or 68 % items those adequate items. It means that those items do not need revise and capable to used as references. Thus, teacher will know the progress of the students correctly.

⁴ Prof. Dr. M. Soenardi Djiwandono, *Tes Bahasa Dalam Pengajaran*, (Bandung: ITB, 1996), p. 140

3. Index of Discrimination

a. Index of discrimination XII Science Class

The table IV.8 showed that 12 out of 25 items have poor index of discrimination. The poor items are number 1, 3, 4, 6, 7, 8, 9, 14, 18, 19, 21, and 23. These items have index of discrimination value between 0,00 - 0,15. And two items, number 16 and 25 have discrimination value -0,05. These items can not be defended. It should be deleted. Heaton stated that 0 = an item which does not discriminate in any way at all. -1 = an item which discriminates in entirely the wrong way⁵.

The items number 2, 5, 11, 13, 15, 17, 20, and 24 have index of discrimination between 0,25 – 0,35. According to Arikunto and Djiwandono, these items classified as satisfactory.

And 3 items have good index of discrimination. And the good items are 10, 12, and 22. These items have index discrimination value between 0,40 – 0,45.

The summarizing of the finding above is 56% or 14 out of 25 items can not be defended and can not be used. These items must be revised. Nevertheless, there are still 44 % items that can be used.

⁵ J.B.Heaton, *Writing English Language Test*, (New York: Longman, 1975), p. 180.

b. Index of discrimination XII Social Class

The finding above reported that 10 out of 25 items have poor index of discrimination. These items are 1, 3, 4, 5, 7, 10, 14, 16, 21, and 25. It have index of discrimination value between 0,00 – 0,20. The item number 4, 7, 16, 21 have index of discrimination very low; 0,00 – 0,07. These items are doubted and do not used. And item number 1 have index of discrimination value -0,07. This item should be discarded.

And 2 items have index of discrimination value 0,27 – 0,33. These items are classified in satisfactory category. These items are 9 and 19. The satisfactory category can be used and not need to be revised. But these items have low discrimination. The higher index of discrimination is better because it is more distinguish the lower and the higher students' ability and vice versa.⁶

The rest items have good index of discrimination. There are 13 items. The items are 2, 6, 8, 11, 12, 13, 15, 17, 18, 20, 22, 23, and 24. It have index of discrimination value between 0,40 – 0,67. These items are proper to used.

In conclusion, there are 15 out of 25 items have good index of discrimination. It means that 60% items that proper and not to be revised.

⁶ Burhan Nurgiyantoro, *Penilaian Dalam Pengajaran Bahasa Dan Sastra*, (Yogyakarta: BPFE Yogyakarta, 2001), p. 140

Finally, the index of discrimination of the test can conclude that both of Science and Social class have different result of index of discrimination. Science class have unacceptable index of discrimination since 44% items can used and the Social class have 60% items that can used, it is means that the Social class has acceptable index of discrimination.

CHAPTER V

CONCLUSION AND SUGGESTION

After analyzing and discussing of the results of this study in the preview chapter, there were the conclusion of the research and suggestion. The conclusion was arranged based on the problem of the study: the content validity, the index of difficulty and the index of discrimination of the teacher-made English try-out test in national examination 2010-2011 for third graders of MAN Sidoarjo.

A. Conclusion

From the finding, it can be concluded the answer from research problems as follows:

1. According to the result, the conclusion of the content validity of teacher-made English try-out test of MAN Sidoarjo is the content validity of the teacher-made English try-out test of MAN Sidoarjo has good content validity since 52% items test covered the indicators of Standard of Graduates Competencies.
 2. The teacher-made English try-out test in national examination 2010-2011 for third graders of MAN Sidoarjo have good index of difficulty, because the Science class have 15 out of 25 or 60 % items and Social class have 17 out of 25 or 68 % items those moderate items. It have index of difficulty 0,30 - 0,70. It means that those items do not need revise and capable to used as references.

3. The index of discrimination of the teacher-made English try-out test in national examination 2010-2011 for third graders of MAN Sidoarjo is different between both of class. The result of Science class showed that 14 out of 25 items or 56% items can not be used because it have index of discrimination poor and wrong. And the Social class has good index of discrimination since 60% items has satisfactory and good criteria.

B. Suggestion

There are some suggestions for the teacher or test-maker who want to construct a test.

1. Because of several indicators not cover in the test, the teacher should modify the test so that it is cover the entire indicator of standard of graduates competence. Moreover, the teacher should refer not only the last national examination but the standard of graduates competence also.
2. Although the index of difficulty of the test is good, there are still several items that too difficult and too easy, so the teacher should revise the items test that are easy and difficult in order there will no items that too easy and too difficult.
3. The teacher should revise the items test that are poor, not discriminate in any way at all and discriminates in wrong way in order to know

distinction the better students and the poorer students because there are still several items that have poor and wrong criteria.

4. Although the finding reported that the test is good, but it has minimal criteria as a good test. So, the teacher should analyze their own test to know whether that test is adequate before it is given to students.

For further researcher: this study focuses to the content validity, index of difficulty and index of discrimination of the teacher-made English try-out test, so there are several criteria of a good test that needed to be researched. Moreover, many institutes or courses construct English try-out test by their self. It is still questioned whether the test is really constructed proper based on the criteria of a good test.

BIBLIOGRAPHY

- Alderson, J. Charles, et al. 1995. *Language Test Construction and Evaluation*. New York: Cambridge University Press.

Arikunto, Suharsimi. 1996. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.

Bachman, Lyle F. et all. 1996. *Language Testing In Practice*. New York: Oxford University Press.

Bogdan, Robert. 2006. *Qualitative Research for Education: An Introduction to Theories and Methods*. New York: Pearson Education.

Brown, H. Douglas. 2004. *Language Assessment: Principles and Classroom Practices*. New York: Pearson Education.

Danim, Sudarwan. 2002. *Menjadi Peneliti Kualitatif*. Bandung: Pustaka Setia

Djiwandono, Soenardi M. Prof. Dr. 1996. *Tes Bahasa Dalam Pengajaran*. Bandung: ITB.

Fulcher, Glenn, 2007. *Language Testing and Assessment*. New York: Routledge.

Heaton, J. B. 1975. *Writing English Language Test*. New York: Longman.

Hughes, Arthur. 1989. *Testing For Language Teacher*. Cambridge: Cambridge University Press.

Islamiyah, Millatul, 2010, ‘*an analysis content validity and item analysis of semester II English final test for tenth grade students of SMAN 3 Sidoarjo*’, thesis S1 Surabaya: perpustakaan IAIN Sunan Ampel.

Madsen, Harold.S. 1983. *Technique in Testing*. New York: Oxford University Press.

Mayangsari, Iffah Mursyidah, 2009, ‘*an analysis of UAS English test of second semester 2008/2009 by teacher-made English test in SMA 2Muhammadiyah Sidoarjo*’, thesis S1 Surabaya: perpustakaan IAIN Sunan Ampel.

Moleong, Lexy J. DR.M.A, 1989. *Metodologi Penelitian Kualitatif*. Bandung: PT Remaja Rosdakarya.

Nurgiyantoro, Burhan, 2001. *Penilaian Dalam Pengajaran Bahasa Dan Sastra*. Yogyakarta: BPFE Yogyakarta.

Rahma, Anita Nur, 2004, '*an analyzed the English summative test items for semester 1 2002-2003 of the first year students of SMU in Surabaya*', thesis S1 Surabaya: perpustakaan UNESA.

Saputro, Sigit Eko, 2006, '*an analyzed the teacher-made test items in UAS semester 1 2005-2006 of the first year students of SMA 2 Magetan*', thesis S1 Surabaya: perpustakaan UNESA.

www.ujiannasional.org [accessed on 1st July 2011]