AN ANALYSIS OF THE TEST ITEMS IN ENGLISH TRY-OUT TEST FOR UN 2010/2011 PUBLISHED BY DIKNAS SURABAYA

THESIS

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



By: ULLIA DWI AGUSTINA NIM: D35207006

ENGLISH EDUCATION DEPARTMENT
FACULTY OF TARBIYAH
STATE INSTITUTE OF ISLAMIC STUDIES
SUNAN AMPEL SURABAYA
2011

AN ANALYSIS OF THE TEST ITEMS IN ENGLISH TRY-OUT TEST FOR UN 2010/2011 PUBLISHED BY DIKNAS SURABAYA

THESIS

Submitted in partial fulfilment of the requirment for the degree of Sarjana Pendidikan Islam (S.Pd.I) in Teaching English

By:

ULLIA DWI AGUSTINA NIM: D35207006

ENGLISH EDUCATION DEPARTMENT
OF TARBIYAH FACULTY
STATE INSTITUTE OF ISLAMIC STUDIES SUNAN AMPEL
SURABAYA
2011

APPROVAL SHEET

This thesis by Ullia Dwi Agustina entitled An Analysis of The Test Items in English Try-Out Test for UN 2010/2011 Published by Diknas Surabaya has been approved by the thesis advisors for further approval by the Board of Examiners.

Surabaya, September 6th, 2011

ADVISOR I

Mokhamad Syaifudin. M.Ed.

NIP. 197310131997031002

ADVISOR II

Sigit Pramono Jati, M. Pd. NIP. 198605182011011005

EXAMINERS APPROVAL SHEET

This thesis by ULLIA DWI AGUSTINA entitled AN ANALYSIS OF THE TEST ITEMS IN ENGLISH TRY-OUT TEST FOR FINAL EXAMINATION (UN) 2010/2011 PUBLISHED BY NATIONAL EDUCATION OF SURABAYA has been examined on September 14, 2011 and approved by the Board of Examiners.

Dean,

Dr. H. Nur Hamim, M. Ag.

NIP. 196203121991031002

The Board of Examiners

Advisor/ Head of Examination,

Mokhamad Syaifudin. M.Ed.

NIP. 197310131997031002

Segretary,

Sigit Pramono Jati, M. Pd.

NIP. 108605182011011005

Examiner I,

Khoirun Niam. Dr. Phil

NIP. 197007251996031004

Examiner II,

Drs. Achmad Zani, MA.

NIP. 197005121995031002

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama

: Ullia Dwi Agustina

NIM

: D35207006

Jurusan

: Pendidikan Bahasa Inggris

Fakultas

: Tarbiyah

cara dan sistematika penulisan yang lazim.

Judul Skripsi : An Analysis of The Test Items in English Try-out Test for UN

2010/2011 Published By Diknas Surabaya

Menyatakan bahwa karya ilmiyah ini adalah benar-benar merupakan hasil karya saya sendiri, dan bukan merupakan pengambil-alihan tulisan atas pemikiran orang lain, kecuali pada bagian tertentu yang digunakan sebagai acuan dengan mengikuti tata

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya, dan jika suatu saat nanti ditemukan ketidakbenaran dalam pernyataan ini, maka saya bersedia mendapatkan sanksi akademis.

Surabaya, 16 Agustus 2011

Ullia Dwi Agustina NIM D35207006

ABSTRACT

Dwi Agustina, Ullia. 2011. An Analysis of the Test Items in English try-out test for UN 2010/2011 Published by Diknas Surabaya. A Thesis. English Department, Faculty of Tarbiyah, IAIN Sunan Ampel, Advisor: Mokhamad Syaifudin, M.Ed.

Key words: Test items and English Tryout Test, Diknas Surabaya.

Testing and teaching are closly interrelated since testing determine the achievement of the objective of education. It means after doing their materials that students taught, the teacher needs to measure the student ability by doing a test. After getting a test, the teacher will know how far their students' master the lesson and what parts are difficult for them.

This thesis is studied about analysis of the test items in English Tryout Test for UN 2010/2011 published by Diknas Surabaya. The writer focuses on problems about the face validity and the content validity of the test items in English try-out test for UN 2010/2011 published by Diknas Surabaya.

This thesis uses descriptive methodology. Because it will describe the quality of the test items in English try-out test for UN 2010/2011published by Diknas Surabaya and another factors that could be determine this test is a good test or not. Also to measure the validity, the writer using a checklist as the instrument.

The result of this study, After analyzing the the face validity, the writer gets the result that English try-out test items for UN 2010/2011 published by DIKNAS Surabaya is proper to conduct to the student. For the content validity, the writer gets the result that the content validity for instance the writer found the result that many test items in test items in the test are appropriate with the SKL. From the explanation about the appropriateness of the test, English try-out test items for UN 2010/2011 published by DIKNAS Surabaya has high content validity because from the whole questions (100%), 82% is appropriate with the SKL.

TABLE OF CONTENTS

COVER	•••••••••••••••••••••••••••••••••••••••	i
APPROVAL SH	IEET	ii
STATEMENT OF GRADUATE		
EXAMINERS A	PPROVAL SHEET	iv
DEDICATION S	SHEET	v
мотто	······································	vi
ABSTRACT		vii
ACKNOWLED	GMENT	viii
TABLE OF CO	NTENTS	ix
LIST OF APPENDICES		хi
LIST OF TABL	ES	xii
CHAPTER I	INTRODUCTION	1
	A. Background of the Study	1
	B. Research Question	5
	C. Objective of the Study	5
	D. Significance of the Study	6
	E. Scope and Limitation	6
	F. Definition of Keyterm	7
CHAPTER II	REVIEW OF RELATED LITERATURE	0
	A 000 4	8 8
	B. The Purposes of a Test	9
		-
	C. Types of a Test 1. Based on the Test-maker	10
	2. Based on the Form of Test Items	10
		13
	3. Based on the Test Purpose	16
	D. Characteristic of a Good Test	19
	1. Validity	19
	a. Face Validity	19
	b. Content Validity	20
	2. Objectivity	
	3. Practicality	22

	4. Economy	22
	E. SKL(Standard of Graduates' Competence)	23
	F. Review of the previous study	23
CHAPTER III	RESEARCH METHOD	25
	A. Research Design	25
	B. The data collection technique	26
	C. Instrument of the Data	26
	D. Data Analysis Technique	27
	1. Analyzing the face validity	27
	2. Analyzing the content validity	28
CHAPTER IV	DATA DISPLAY AND DISCUSSION	29
	A. Result	29
	1. Face Validity	
	2. Content Validity	
	B. Discussion	52
CHAPTER V	CONCLUSION AND SUGGESTION	54
	A. Conclusion	
	B. Suggestion	

CHAPTER I

INTRODUCTION

A. Background of the Study

Teaching English is teaching the skills which are emphasized in listening, speaking, reading and writing. To know whether the students have mastered English or not, the teacher need to use the test.

Heaton states that both teaching and testing are closely interrelated it is virtually impossible to work in either field without being constantly concerned with the other. It means after doing their materials that students taught, the teacher need to measure the student ability by doing a test.

Test is important to measure students' mastery from the lesson given. Arikunto states that test is any series of questions or exercise or other means of measuring skill, knowledge, intellegence, capacities of aptitudes or an individual or group.² In a field of study, student will study hard if they know that the end of their study, they will have a test. After getting a test, the teacher will know how far their students' master the lesson and what parts are difficult for them.

A well-constructed classroom test will provide the students with an opportunity to show their ability to recognize and produce forms of the language. In this way, a

¹ J.B.Heaton, Writing English Language Test, (New York:Longman, 1975), page 1

² Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, (Jakarta:Bumi Aksara, 1984), page 29

good test can be used as a valuable teaching device. If details of his performance are given as soon as possible after the test, the students should be able to learn from his errors and consolidate the pattern taught. According to Heaton, a good classroom test will also help to locate the precise areas of the difficulty encountered by the class or by the individual students. Therefore, it is necessary for teacher to diagnose their students' weakness and difficulties. By knowing the students' weakness, teacher will know the cause and it will be easier for them to solve the learning problem.

Based on the function of testing above, there are two kinds of test according to the test maker. Those are standardized-test and teacher-made test. Harris states that standardized test is designed with thousands subjects through the nation or the world and prepared by a testing specialist.³ Standardized test is made by the professional test maker and tested several time. Standardized test is do by all the students with the same guidlines and also in the same time. There are the examples of standardized test are UN, SPMB, Try-Out, etc.

Different from standardized-test, teacher-made test is a test made by the teacher her/his self or group of teacher without tried on first, revised, and analysed. Arikunto states that teacher made test is a test made by the teacher him/her self or a group of theacher using untried out, unanalysed and unrevised test items.⁴ The example of teacher made test are middle test and final test(UAS).

⁴ Suharsimi Arikunto, op.cit, page 144

³ David Haris P, Testing English as a Second Language, (New York: Mc Graw-Hill, Inc, 1969), page 2.

The explanation above shows that test is the important things to measure student achievements. Therefore, the students must be able to pass the test when they want to finish their study and they also must be able to manage their time well before facing the UN test. In the highest grade of school (for examples 6th grade of elementary school, 9th grade of junior high school and 12th grade of senior high school), students have to pass the final eaxamination called UN.

In the decree of national education minister number 78 in 2008 (part 1) explains that the final examination which is usually called by UN is a kind of measurement for the students from a basic level till intermediate level. (part 2) the purpose of final examination is to measure the students' achievement in several particular lessons in the national level. (part 3) the result of the final examination can be used to map the quality of education program, to select the student who will continue to the next grade, to determaine the students' graduation, to construct and assist units of education in increasing the quality of education.⁵

Realizing the important of final examination (UN), the National Education of Surabaya gives a try-out test. With the try-out test, the student to be able to prepare in doing the final examination (UN) test. The try-out test is needs to know how far the students are ready to do UN. To face the final examination, students will not be nervous when they do the real test. Therefore, for this problem Diknas Surabaya create try-out test which is based on the current curriculum. There are several tryout

⁵ Salinan Peraturan Pemerintah Pendidikan Nasional no 78 tahun 2008, (http://www.dinpendikpkp.go.id/file/permen/smk.pdf, assed on September 15, 2011)

was done in 2010/2011 published by Diknas Surabaya, there are two code (A&B). But the writer choose this one code (B). Because she gets only one tryout test from Dinas Surabaya and the other code is the same test and just the arrangement of the number are different.

Based on the try-out test which is published by Diknas Surabaya, the writer will analyze the content validity and the face validity. She wants to know the try-out test which is published by Diknas Surabaya have good content validity and face validity or not.

According to Heaton, there are four types of validity, namely face validity, content validity, construct validity and empirical validity.⁶ From that types of validity, in this study, the writer focus only on the content validity and the face validity. Because with content validity, the test does represent all the content stated in the subject or not. Beside, with the face validity, According to Hughes, the test said to have face validity if it looks if it measure what it supposed to measure. A test which does not have face validity may not be accepted by the candidates, teachers, education authorithies or employers.⁷ This is the reason why the writer does not focus on all validity it is just because she need more time to focus on all validity and she needs the time any longer to analyze another validity.

Heaton also states, in face validity, in fact, if a test items look right to other testers, teachers, moderators and testees, it can be described as having at least face

⁶ J.B Heaton, op cit, page 135

⁷ Arthur Hughes, *Teasting for Langage Teacher*, (UK:Cambridge university press, 1989), page 33

validity.⁸ From the explanation before, face validity is the extent of a test to measures the performances of a test. It can be concluding copying, typing, instruction, title, organization, etc. Face validity concerns about the test looks valid to the examainees who take it. From the explanation before, it can be concluded that face validity is the way the test looks and the way the test is organized.

For content validity, According to Hughes, a test is said to have content validity if its content constitutes a representative sample of the langage skills, structures, etc.9 The test is said content valid if the test had been appropriate with the lesson.

B. Research Question

Based on the background of the study above, the writer will set out the statement problems, they are:

- 1. How is the face validity of the English try-out test for UN 2010/2011 published by Diknas Surabaya?
- 2. How is the content validity of the English try-out test for UN 2010/2011 published by Diknas Surabaya?

C. Objective of the Study

Based on the research problems, the objectives of the study are stated as follows:

⁸ *Ibid*, page 153
⁹ Arthur Hughes, *op cit*, page 26

- 1. To find out the face validity of English try-out test for UN 2010/2011 published by Diknas Surabaya.
- To find out the content validity of the English try-out test for UN 2010/2011 published by Diknas Surabaya.

D. Significance of the Study

This study is expected to be useful to the teaching learning English, including for the teachers and the test makers:

- For the teachers: the writer expects this analysis will be the information for teacher in increasing their knowledge about the preparation in doing try-out test and the teacher can motivate the student about the materials that often use in English try-out.
- 2. For the test maker: the writer expects with this analysis the test makers can construct the test better than ever.

E. Scope and Limitation

The scope and limitation of this study is the English try-out test for UN 2010/2011 published by Diknas Surabaya. In this study the writer will focus on the B code of tryout test and the content validity also the face validity. After analyzing the content validity the writer match it with the SKL(Standard of Graduates' Competence) listed in the decree of national education minister number 78 in 2008.



F. Definition of KeyTerms

To avoid misunderstanding and misinterpretation to the terms used in this study, the following definitions are given:

- 1. Try-out test is a test usually made based on the current curriculum by competent courses or institutions which is published by DIKNAS Surabaya.
- 2. UN test : UN is a test used to determine whether the students are able to graduate from the school or not.
- 3. Standardized-test: Standardized-test is a test made according to the content and the general goal for all school in the country and it was related with general knowledge and capability.¹⁰
- 4. Face validity : Face validity is the way the test items look right to other testers, teachers, moderators and testees. 11
- 5. Content validity: A test is said to have content validity if its content constitutes a representative sample of the language skills, structures, etc. 12

¹⁰ Arikunto suharsimi, op cit, page: 144

¹¹ Arthur Hughes, *Teasting for Langage Teacher*, (UK:Cambridge university press, 1989), page 153 ¹² *Ibid*, page 26

CHAPTER II

REVIEW OF RELATED LITERATURE

This chapter discusses some literature which is closely related to the questions of this research. This literature will focuse on: test, the purpose of the test, types of a test, characteristic of a good test, SKL (Standard of Graduates' Competence), and review of the previous study. Thus, each of these main parts is devided into several sub part.

A. Test

A test is one of type of measurement used to measure the individual's behavior change as the goal of instruction. By giving a test to the students and evaluating the result, the teacher will know exactly the students' changes on their behavior. Arikunto states that a test is any series of questions of exercise or other in measuring skill, knowledge, intellegence, capscities of aptitudes or an individual group.

On the other hand it is unquestionable that there are many people are often confused by the terms evaluation and measurement. They often think that the terms are similar, but they are actually not that way. According to Gronlund, evaluation is used as a collective term for those appraisal methods which does not depend on the

¹ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, (Jakarta: Bumi Aksara, 1993) page 29

"measurement".² By the statement, we can see that evaluation is more comprehensive and inclusive then measurement. Nurgiyantoro also states that measurement, in this case is only an evaluation's part.³

Besides, Nurkanca and Sumartana said that evaluation refers to the act of process to determine the value of something.⁴ They explain that educational evaluation is a process of determaining the value something in the educational field.

B. The Purposes of a Test

Test is used to measure students' mastering with the subject given. Some experts mention the other purpose of test. According to Nurkanca and Sumartana, a test has many purposes. First, is to know how far the result of a progammer applied whether it has reached its goal or not. Second, is to see whether the materials should be re-taught or not. Third, is to get some information about the students' weakness and difficulties in learning about the given materials. Fourth, is to determine the students' achievement and to allow them going through to the grade. Fifth, is to select and group students based on their achievement.

² Norman.E.Gronlund, *Measurement and Evalution in Teaching*, (New York: mcmillan publishing Co.Inc, 1976) page 6

³ Burhan Nurgiyantoro, Penilaian Dalam Pengajaran Bahasa dan Sastra, (Yogyakarta: BPFE, 1988) page 5

⁴ Nurkancana, Wayan dan Sumartana, Evaluasi pendidikan, (Surabaya:usaha nasional, 1986) page 1

Test is to measure students' mastering with the suject given. Some experts mention the other purposes of the test. According to Nurgiyantoro the aims of a test are:⁵

- 1. To find out how far the purpose of educatinal can be achieved.
- 2. To give an objectivity in evaluating the result of teaching learning process.
- 3. To find out the student ability in certain fields or topics.
- 4. To determine whether or not a student as able to move to higher grade or graduated.
- 5. To give feed back for teaching and learning process.

From the statement above we can concludes that the main purpose of test is to measure the students' ability, to find the effectiveness of a program applied, determine students' weakness and find the program solving.

C. Types of a Test

1. Based on the Test-maker

There are two kinds of a test based on the test maker. They are standardized test and teacher made test.

a. Standardized test

Standardized test is a test made by professional test makers, administered and scored in a consistent manner. Harris stated that standardized test is a test designed to

⁵ Burhan Nurgiyantoro, op cit, page 14

be used with thousands and sometimes hundred thousands of subjects trhoughout the nation or the world, and prepared by team of testing specialists.⁶

There are some characteristics of standardized test. According to Arikunto a standardized test is a test made according to the content and the general goal for the whole country and it is in relation with the general knowledge and capability. It is developed by professors, reviewers and editors of test items. The items of test are tried out, analyzed and revised before being used for a test.

Dealing with Arikunto, there are four purposes of standardized test. The first is to compare students' achievement with individual's character or group's character, to compare students' achievement in an instructional in an group or individual, to compare student achievement in several schools and to deternaine students' development in a certain period of time.

There are some examples of standardized test such as: Final Examination, TOEFL, IELTS, try-out test, etc. Tryout test is a test hold before UN. Try-out test is consists of two words those are "try" and "out" which each of them has the main purposes. "try" has two purposes. First, is to measure how far the students' readiness about the matter that will be tasted in UN and second is to measure the test it self (it means that the test is not just anything taken but it must be constructed based on the

⁶ David P Harris, Testing English as a Second Language. (New York: Mc Graw-Hill.Inc, 1969) page 2

¹ Ibid, page 138

curriculum). "out" also has two purposes first is to encourage the students in facing UN and second is to make the students get used to face a situation like in UN.

b. Teacher made test

Teacher made test usually prepared, administered, and scored by the teacher her/himself. According to Harris, teacher made test are generally prepared, administered, and scored by one teacher. It usually untried out first, unanalyzed and unrevised before being used for test. Arikunto also states that teacher made test is a test made by the teacher him/herself or group of teachers and untried out, un analyzed and unrevised the test items.⁸

There are the characteristics of teacher made test. according Arikunto, teacher made test is constructed according to the content or special goal given by the teacher and it is certain with capability and knowledge.⁹

There are three purposes of teacher made test. According Arikunto, the first is to measure achievement withthe instructional during certain time, the second is to whether a purpose of instructional has achieved or not nad the third is to get the students' scores.¹⁰

⁸ David P Harris, op cit, page 13

⁹ Suharsimi Arikunto, op cit, page 144

¹⁰ *Ibid*, page 139

2. Based on the Form of Test Items

a. Subjective Test

Subjective test is a test that needs the student to answer with their opinion, not only to choose the option or something available. It is usually called by essay test. According to Nurgiyantoro, essay test will give students freedom in arranging and expressing of their in relatively in limited scope. If Arikunto also states that subjective test is a test has strength and weakness. The strength of subjective test is easy to prepare and give an opportunity to the students to express their idea with their own words but it gives not opportunity to speculate or luckiness. In

The weakness of subjective test is fewer representative whole subject of instruction because the question is limited.it will be difficult to correct since it needs an individual consideration and the score will be influenced by scores' objectivity.

b. Objective Test

Objective test is a test that encourages the student to choose the correct answer or give short answer. Nurgiyantoro defines objective test as a short answer test. 13 According to Arikunto, objective test has strength and weakness. The strength of an objective test is more representatives the whole subject instructional and it is does not need scores' subjectivity in correcting. It is easy to correct and it can be given to other

¹¹ Burhan Nurgiyantoro, op cit, page 71

¹² Suharsimi Arikunto, op cit, page 155

¹³ Burhan Nurgiyantoro, op cit, page 75

people to score it.¹⁴ Objective test is a term refers to scoring of the test that can be describe as objective. It means that the testee will get the same score, no matter which examiner which marks his test, since these test have usually have one answer. Moreover, the objective test is the test that can be scored objectively. The number of item is usually more than the essay test, the testee have to to weigh up to carefully all the alternatives and select the best one.

The weakness of objective test is difficult to construct. The test only measure the students in memorizes. The result of the test sometimes is influenced by luckiness since the students can guess the answer or chest each other.

There are many varieties of new types. But four kind are most common use, true or false, multiple choice, completion, matching. In this case the writer will focus only on multiple choice test.

A multiple choice test is a kind of objective test. which asked the students to choose one of the alternative available as a correct answer. According to Nurkanca and Sumartana that a multiple choice items is an item which consists of incomplete statement or stem. Therefore, multiple choice test consist of stem, which present a problem situation and several options, which provide possible solution to the problem. Not only the correct answer but also several wrong answer or distracters in which to distract those students who are uncertain of the answer found in the option answer.

14 Suharsimi Arikunto, op cit, page 156

One of the basic characteristics of multiple choice is the distracters. The distracters called good distracters if it will attract more the students from the lower group than the upper group and when item distracters attract the students from the upper group than the lower group, it is not good distracters.

There are some guidlines how to construct an item for multiple choice test, as below:

- Each multiple choice item should have only one answer. This answer
 must be absolutely correct unless the instruction specifies choosing the
 best option.
- 2. Only one feature at a time should be tested. It has long been standard practice to test only one feature at a time. It usually less confusing for the testee and it help to reinforce a particular teaching point.
- 3. Each option should be grammatically correct when placed in the stem, except of course in the case of specific grammar test items.
- 4. All multiple choice items should be at level appropriate to the linguistic ability of the testees. The cotext it self, should be at a lower level than the actual problem which the item is testing.
- 5. Multiple choice items should be as brief and as possible (thought it is often described to provide short contexts for grammar items)
- 6. In many tests, items are generally arranged in rough order of increasing difficulty. It is generally be considered important to have one or two

simple items to "lead in" the testees, especially if they are not too familiar with the kind of test being administered.

3. Based on the Test Purpose

Based on the purpose of a test, the writer will explain about the kind of achievement test.

a. The Achievement test

Achievement test uses to measure students' achievement in order that the students will know their progress of their knowledge through this kind of test. Nurgiyantoro states that there are four types of achievement test, they are pretest, diagnostig test, formative test and summative test. 15

b. Pretest

Pretest is a test or measurement taken before service or activities begin. It is compared with the results of a posttest to show evidence of the effects of the learning; they will know the result of the teaching by comparing the students' scores in pretest and posttest. According to Nurgiyantoro, there are two kinds of pretest. They are entry behavior test and placement test. ¹⁶

(a). Entry Behavior test

¹⁵ Burhan Nurgiyantoro, op cit, page 65

¹⁶ Burhan Nurgiyantoro, op cit, page 62

The students study lessons they taught and after that they take a test. This test aim is to know whether the students reach specific ability or skill as in the requirment. For example, a student in English department may take the speaking II if he has pass the speaking I.

(b). Placement Test

The test will evaluate the skill level and the palce in the right class. By giving this test, the teacher will be easier to teach their student with the same ability. According to Gronlund, the goal of the placement evaluation is to determine in the instructional squence and the made of instruction that is most likely to provide optimum achievement for each pupil.¹⁷

(c). Diagnostic Test

Diagnostic test concern with the pupil's difficulties during instructional squence. The purpose of diagnostic test is to ascertain students' strengths, weaknesses, knowledge, and skill. A diagnostic test holds assignment from a wide range of fields. This kind of preliminary knowledge is vital when devides the students into groups. A diagnostic test helps the teacher in adjusting the curriculum to meet the problem solving of the student difficulties. According to Gronlund, the main aim of diagnostic test is to determine the causes of learning problem and to formulate a plan for remedial action. ¹⁸

¹⁷ Norman E Gronlund, Masurement and evaluation in Teaching, (Newyork: 1981), page 17

¹⁸ Ibid, page 18

(d). Formative test

Formative test gives after the teachers finish the unit of a lessons. It uses to gives a feed back of students' understanding and to know students' development. Gronlund states, the function of formative test is to determine learning progress, provide feed back for reinforcement for the students. If their grades are good, they will keep their study based on their way. Giving this test will encourage students to study hard in order to get a good mark as a result of that test. in spite of this, teacher also will be able to know the effectiveness of their method of teaching.

(e). Summative test

Summative test is the formal testing of what has been learned in order to produce marks or grades which may be used for report of various types. The main purpose of summative test is not only for grading or certifying pupil's achievement, but also provides the information for the judging the course objectives and effectiveness of instructional. Johnson and Johnson states that summative test is conducted at the end of an instructional unit of the semester to judge the final quality and quantity of student achievemnt and or the success of instructional program.¹⁹

¹⁹ David W, Roger T Johnson and Johnson, *Meaningful Assessment: a Meaningable and Cooperative Process*, (USA:Allyn and Bacon, 2002), page 7

D. Characteristic of a Good Test

A testing is an important part in language learning process, it is essential to know the characteristic of a good test so that the test administered can represent the degree of students' mastery over the language teaching materials have been taught.

According to Harris, that a test can be qualified one if it given the fulfills criteria. They are validity, reliability, and practicality.²⁰ According to gronlund, validity and reliability are the most important thing related to the test.²¹ In this study the writer will focus on validity.

1. Validity

Validity is an important thing that must be considered before constructing a test. Heaton states that the validity of a test is the extent to which is measures what it is supposed to measure and nothing else. There are four type of validity. According to Heaton, there are four types of validity there are face validity, content validity, construct validity and empirical validity.²² In this study the writer will focus only on face validity and content validity.

a. Face Validity

According to Heaton, in this type of validity, in fact, is often referred to as face validity:if a test items look right to other testers, teachers, moderators, and

²⁰ David P Harris, op cit, page 13

²¹ Norman.E.Gronlund, op cit, page 130

²² J.B Heaton, Writing English Language Test, (London: Longman Group, 1975) page 153

testees. It can be described as having at least face validity. Therefore, face validity is the extent of a test measures the performance of a test. it can be concluding copying, typing, instruction, title, organization, etc.²³

Based on the statement above, the writer concludes that the face validity of a test refer to:

- a. The way the test "looks" based on the testees, the administrator, and the educators' point of views.
- b. The way the test is recognized. It consist of types of test and test directions.

Therefore, if the test does not have face validity, the students who take it may not perform a reflection of their truly abilities. The face validity is also need to funtion effectively in practical situation. It means that we can evaluate the face validity of a test by comparing how the test "looks" with the criteria of how to construc a good test.

b. Content Validity

Content validity is the extend to which a test measures the representative sample of the subject matter content. Hughes said that a test is said have content validity if its content constitutes a representative sample of the language skills,

-

²³ Ibid, (1988) page 153

structure, etc.²⁴ It can be said that the test had high content validity if the test items cover the materials that are stated in curriculum.

To measure whether the test has content validity or not, the writer compares the test materials with the materials stated in the subject matter or curriculum, it can be said that the test has content validity.

On the other hand, if the test materials are not suitable with the subject matter curriculum, it means that the test does not have the content validity and it is not valid. According to Gronlund, there are three ways to make test had high content validity. The first is identifying the subject matter topics and behavior outcomes to measures. The second is building a table of specification, which specifies the sample of items to be used, and the third is constructing a test that closely fits the table of specifications.²⁵

2. Objectivity

A test is said to be objective if it is free from personal biases in interpreting its scope as well as in scoring the responses. According to Arikunto, a test had high objectivity if there are no subjective factors in the process administering. Objectivity of a test can be increased by using more objective types test items and the answers are

²⁴ Arthur Hughes, *Teasting for Langage Teacher*, (UK:Cambridge university press, 1989), page 26

²⁵ Norman.E.Gronlund, op cit, page 132

scored according to model answers provided.²⁶ Arikunto adds that there are two factors that influence the objectivity of a test they are the form of a test and the test scorer.²⁷

3. Practicality

A test is called as practical test if it is easy to administered(does not require many equipments and give freedom to the students to do the easier part), easy to score, is completed with clear instructions. Arikunto stated that practicality of a test deals with a level of difficulties in admintering the test it self.²⁸

4. Economy

According to Arikunto the term economy in a test deals with the amount money, time and energy that a test taker spends to take a test. from the definition before, it can be concluded that administering a test requires considering manyfactors, such as money, time and energy that the test takers might spent.²⁹

E. SKL(Standard of Graduates' Competence)

²⁶ Suharsimi Arikunto, op cit, page 56

²⁷ Ibid, page 56

²⁸ Ibid, page 58

²⁹ Ibid, page 58

SKL (Standard of Graduates' Competence) is an intersection of competence Based Curriculum which takes several objectives study from the entire curriculums. While Competence Based Curriculum (CBC) is curriculum which is designed to help the users to develop English communicative competence as life skills, the skills to make learner survive in the modern communications where English is used.

According to Agustien the 2006 English standard competency is the improvement of the 1994 curriculum where the students asked to be active in the class. It is different from the latest curriculum which lets the student only listen what the teachers told even if the basic of this curriculum is to make students able to speak English actively.

In this curriculum also covers the competency of reading, listening, writing and speaking. With those four skills can make the students able to communicate and to create discourse in English.

F. Review of the Previous Study

In this part, the writer presents the result of the researcher concerning with the writer research. Dealing with this study, there are three studies taken before. Those are:

1. The thesis was done by Repelita Wardani (UNESA, 1995), in her thesis, she analyzed English Summative test written by English teachers of senior high school. In her thesis, she focused on the objective test. she found that the test

had high validity, adequate reliability, acceptable facility value and effectiveness of the distracter. Nevertheless, the test did not adequate discrimination index.

- 2. The thesis was done by Nurul Koiriyah (UNESA, 1996), in her thesis, she analyzed reading section of the English items of UAN 2003/2004. In her thesis, she found that the test had high content validity, acceptable reliability, did not acceptable index of difficulty, had poor discrimination index and had effective distracters.
- 3. The thesis was done by Moch. Solihan (UNESA, 2000), in his thesis, he analyzed the teacher made English test items in formative semester I 2008/2009 of the first year students of SMAN 22 Surabaya. in his thesis, he found that the test had high content validity and adequate reliability but it did not acceptable index of difficulty, and the index of discriminatin in multiple-choice test was not adequate.

CHAPTER III

RESEARCH METHOD

A. Research Design

Dealing with the research problem and the objective stated in the previous chapter, in this study the writer will use descriptive methodology. This descriptive study is designed to obtain information concerning particular issues and then describe them. Arikunto states that descriptive research is not meant to test a certain hypothesis, but it only describes the phenomena, situation and condition that occur during the study. Descriptive research is to present a board range of activities that have in common the purpose of describing situation or phenomena.

In this study will use descriptive methodology because it will describe the quality of English try-out test published by DIKNAS Surabaya and another factors that could be determain this test as a good test or not. Also thewriter will focus only on face validity and content validity.

¹ Suharsimi Arikunto, Dasar-Dasar Evaluasi Pendidikan, (Jakarta, 1986) page 212

B. The data collection technique

The technique for collecting the data in this study, the writer will use the study documentation in collecting the data. Arikunto states that documentation is a method to get about anything on the form of notes, transcripts, magazines, book, etc.² And than with documentation, the writer will collect the data through these following steps:

- a. Finding the English try-out test for UN 2010/2011 published by DIKNAS Surabaya.
- b. Finding the SKL(Standard of Graduates' Competence) in the decree of national education minister number 78 in 2008.

C. Instrument of the Data

According to Arikunto, research instrument is tools or facilities that are using by the researcher in order to collect data. So that make the job easier, complete and systematic.³

This study used **checklist** to measure the validity of the collected data and to know the test analyzed was valid or not.

² Suharsimi Arikunto, *Prosedur Penelitian, (*Jakarta: PT.Rineka Cipta, 2002) page 234

³ Suharsimi Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktek*, (Jakarta: PT Adi Mahasatya, 2006), page 149

D. Data Analysis Technique

To get the data for this study, the writer initially analyzed the data based on the research problem mentioned in the previous chapter. The analyzed data starts from the first question and the second question will be done as follows.

1. Analyzing the face validity

To begins analyze the face validity of the English try-out test for UN 2010/2011 published by DIKNAS Surabaya, the writer used the theory from Hopkin and Antes. As taken from Adin Iswahyudi, the analysis of the face validity is made in order to evaluate the test "looks", test instructions and the items types. In analyzing the test "looks", it can be seen from the test cover, the size of letter and the paper used. The test instructions are also considered whether the general instructions of the test are presented, how specific instructions of the test are and how the instruction for going on to next section or the next page are presented.⁴

After getting the theory in analyzing the face validity, the writer takes the following steps:

- 1. Classify the test items based on criteria.
- 2. Put the item that have been classified above into a table.

⁴ Charles D Hopkin & Richard L Antes, Classroom Measurement and Evaluation (USA: Peacock Publishers Inc, 1990), page:

3. Conclude the result on the analysis made.

Then the results of the analyses presented in the form of table along with the explanation on chapter 4(see appendix.....)

After analyzing the face validity of the test, the writer gets some result for instance the cover of the test uses blue color, can be read easily and interesting. Next, going to the second question, the writer analyzed the content validity.

2. Analyzing the content validity

To analyze the content validity of English try-out test for UN 2010/2011 published by DIKNAS Surabaya, the writer also takes the following steps:

- 1. Making list of the content English curriculum based on SKL. Putting them into the table.
- 2. Placing each of the test items in the appropriate place with the SKL to identify whether or not the SKL covered by English try-out test for UN 2010/2011 published by DIKNAS Surabaya.
- 3. Counting the precentage of the test items of every language aspects.
- 4. Concluding the result of analysis.(see appendix....)

After analyzing the content validity for instance the writer founds the results that many test items in test items in the test are not appropriate with the SKL. The test is consists of fifty questions so we can say that every test items contains 2%.

CHAPTER IV

FINDING AND DISCUSSION

A. Result

In this part the writer shows the result of her research in face validity and content validity of the test item in English try-out test for the final examination(UN) 2010/2011 published by National Education of Surabaya.

1. Face Validity

To show the result of the face validity of the English try-out test for the final examination(UN) 2010/2011 published by National Education of Surabaya, the first step is classifying the matter to make the test it self. The second step is classifying the test items based on the criteria and then put the items that has been classified into the table and the last is concluding the result.

The test is printed in legal paper. It consists of 8 pieces of paper. The cover of the test is blue color. This test is published by DIKNAS Surabaya and made by the educational agency department of Surabaya named MKKS SMA NEGERI SURABAYA. The first page is used for a cover of this test. The second page is used for the tittle of the test, date and time when the test must be conducted, and several instructions to do the test. The third until seven page is used for listening section. The examples of listening section given on the third page. Question number 1-10 are used

for dialogues and question number number 11-15 are used for monologue. In the reading section, in this test there are 12 texts which are put on page 8-18. This test consists of listening and reading section which consist of 50 items, 15 items for listening section and 35 items for reading section.

Tabel 1 below shows the result of the analysis of the face validity of the test.

Tabel 1

No.	Aspect of test and questions	Explanation
1.	Test looks 1) How is the cover of test?	- The test cover of this test is used blue color, can be read easily and interesting.
	2) How is the letter used in the test?	- The size of the letter in this test used 12. Can be read easily.
	3) How is the test layout?	- The test has good layout.
	4) How is the size of test paper used?	- The paper in this test is used legal paper.

2.	The direction	
	1) How is the general instruction of the test?	- The general instruction of this test are understanable.
	2) How is the specific instrument of the test?	- This test has a specific instrument (tape recorder).
	3) How is the instruction for going on next section of the next page?	- This test has no instruction going to the next section/ending.
3.	Test items types	
	 How are the types of the test items choosen? How are the text presented? 	 In this test there only one type test, Multiple choice test type. This test has quite well presented in the layout or arrangement.

2. Content Validity

A test has a high content validity if the test covers the materials stated in the SKL (Standard of Graduates' Competence). To know whether the English try-out test for the final examination(UN) 2010/2011 published by National Education of Surabaya

has high content validity or not, table specification from the SKL is matched with the material in the test. by placing each item number from the test in an appropriate place of the indicators from the SKL, the representative samples can be identified. In addition, the result shows whether the English try-out test for the UN 2010/2011 published by Diknas Surabaya has high content validity or not.

Table 2, for the first SKL is about listening skill, contains five columns, those are: Standar Kompetensi Lulusan, coloumn which consists of the content of listening skill, Kemampuan Yang Diuji coloumn which consists of the content of each language skill, test item coloumn which consists of the test items appropriate with the objective, relevance coloumn which consists of the relevances between the test items with the SKL and percentage coloumn which consist of the percentages of the items test number with the objective.

Table 3, for the second SKL is about reading skill, contains five coloumns, those are: Standar Kompetensi Lulusan, coloumn which consists of the content of listening skill, Kemampuan Yang Diuji coloumn which consists of the contents of each language skill, test item coloumn which consists of the test items appropriate with the objective, relevance coloumn which consists of the relevance between the test items with the SKL and percentage coloumn which consists of the percentages of the test items number with the objective.

Tabel. 2

Standar	Kemampuan	Test items	Relevance	Percentage
Kompetensi	yang diuji			
Lulusan				
1. Listening	1) Menentuk	- Question	1. this question is	
Memahami	an	number 1:	relevant with the	
makna teks lisan	gambaran	Man: let's	SKL(menentukan	
berbentuk teks	umum,	discuss this	informasi rinci dari	
fungsional	informasi	tomorrow.how	percakapan)(see	
pendek,	tertentu	about nine?	appendix)	
percakapan dan	dan	=>why cant the		
teks monolog	rinci	woman meet		6%
sederhana		tomorrow?		
berbentuk naratif				
(narrative,		- Question	2. this question is	
recount, news		number 2:	relevant with the	
items) dan		Woman: I am	SKL(menentukan	
deskriptif (report,		going to give my	informasi tertentu	
descriptive,		mother this	dari	
explanation)		beautiful bracelet	percakapan)(see	
dalam kehidupan		for her	appendix)	

sehari-hari.		birthday		
		=>what does the		
		man suggest the		
		woman to do?		
		- Question		
		number 3:		
		Man: you look	3. this question is	
		so tired, fanny.	relevant with the	
		Woman: yes, its	SKL(menentukan	
		been tiring after	informasi rinci dari	
		long flight from	percakapan)(see	
		london	appendix)	
		=>what does the		
		man suggest the		
		woman to do?		
	2) Memberik	- Question	4. this question is	
	an respon	number 4:	relevant with the	
	yang	Did you have a	SKL(kepuasan)(see	
	tepat	good trip?	appendix)	
	terhadap		6. this question is	

Т	1		,	T
	percakapa	- Question	relevant with the	10%
	n yg berisi	number 6:	SKL(undangan)(se	
	ungkapan	Man: have you	e appendix)	
	simpati	received the		
	•••••	invitation?		
		- Question	7. this question is	
		number 7:	relevant with the	
		A: will have a	SKL(simpati)(see	
		full month	appendix)	
		vacation next	-pponum,	
		semester		
		- Question	8. this question is	
		number 8:	relevant with the	
		A:Do you still	SKL(kepuasan)(see	
		flight to Jakarta		
		by Parkit	appendix)	
		Air?		
				ı

3) Menentuk an gambar yang tepat sesuai dengan	- Question number 9: A: did you go to the play last night?(see appendix) Monologue Text W: Your mobile beside delighted to know what we bring	9. this question is relevant with the SKL(kepuasan)(see appendix)	
informasi	- Question number 10: =>What picture goes with the conversation?	10. this question is relevant with the SKL(menentukan gambar yg tepat daripercakapan)(se	2%

4	4) Menentuk	Monologue Text	e appendix)	
	an gambar	Petter, when you		
	yang	read this letter I		
	sesuai	am already on		
	dengan	my way to		
	teks	ABC	11. this question is	
	monolog	- Question	relevant with the	
		number 11:	SKL(menentukan	
		Which the 5	gambar yg sesuai dg	2%
		pictures mathces	teks monolog)(see	
		with Mr.James	appendix)	
		Smith?	appendix)	
5	6) Menentuk	Monologue Text		
	an	Hi, Winblue		
	gambaran	the roof of		
	umum,	Bobulact Public		
	informasi	manifancerminal		
	tertentu,	in Bogor		
	dan			

			
	number 14:	appendix)	
	=>What is		4%
	monologue		
	about?		
	- Question		
	number 15:	15. this question is	
	=>Why did the	relevant with the	
	student laugh?	SKL(menentukan	
		informasi tertentu	
		dari teks	
		monolog)(see	
		appendix)	
	- Question		
	number 5:	5. this question is	
	Man :	not relevant or	2%
	did you read the	appropriate with the	
		SKL(Congratulation	
	matter police)(see appendix)	
	news this		
	morning?		
L			

From the result above, in listening section, most of the items in the English tryout test for UN 2010/2011 published by Diknas Surabaya are appropriate with the SKL. For example question number two:

2. Woman : I am going to give my mother this beautiful bracelet for her birthday

Man : Just like that?any plan package? Don't you need any rabbit first and

put it in ribbon on it?

Question : what does the man suggest the woman to do?

Answer : (C), buy her mother a different gift.

This question is relevant with the SKL "Menentukan gambaran umum, informasi tertentu dan informasi rinci dari sebuah percakapan" because this question asks about information related with the dialogue.

However, there was only one items which is not appropriate with the SKL. For number five:

Man : did you read the matter police news this morning?

Woman : No, what's that?

Man : you are the first story telling contest. Congratulation!

Woman :.....?

Answer : (D) Really, I can't trust my ears.

➤ This question is not relevant or appropriate with the SKL. Since it is kind of congratulation expression and does not cover in the SKL.

Tabel. 3

Standar Kompetensi Lulusan	Kemampuan yang diuji	Text/Item Test	Relevance	Percentage
2. Reading	1. Menentukan	1. text 5	Text 5 is kind of	
Memahami nuansa makna di dalam teks tertulis seperti teks fungsional pendek dan esai berbentuk	gambaran umum, informasi tertentu, informasi rinci tersurat, informasi rinci tersirat	1. text 5 This parking garage is closed for repairs. You can park 24 hours a day in the garage on Second Avenue between a.m and 9 p.m. or in the lot at the bus station across the street 24 hours a day Question Number 30:	Text 5 is kind of announcement or notice of parking garage is closed for repair and relevant with the SKL.	
naratif (narrative, recount, news item), deskriptif (report,		Why is the garage closed? Question Number 31: What is suggested to the reader? Question Number 32: What is the type of above text?	menentukan informasi rinci tersirat. QN 31:menentukan informasi rinci tersurat.	6%

descriptive,				QN	
explanation)				32:menentukan	
dan				gambaran umum.	
argumentatif					
(exposition,			Text 12	Text 12 is kind of	
discussion)			No trespassing,,private	announcement to	
dalam			propertydo not enter	the office boy to	
konteks			Question Number 50:	not trespassing	
			What should reader	the room which	
kehidupan			do?	inside there are	2%
sehari-hari.				office document	
				and nobody can	
				enter.	
				QN	
				50:menentukan	
				gambaran umum.	
	2. Mener	ntukan			
	gamba	ran umum,			
	inform	asi tertentu,			
	makna	kata, tujuan			

 	**************************************			1.
			tertentu.	
	Managari			
4.	Menentukan			
	gambaran umum,			
	informasi tertentu,			
	informasi rinci			
	tersurat, pikiran			
	utama paragraf			
_				
5.	Menentukan			
	gambaran umum,			
	informasi tertentu,			
	informasi rinci			
	tersurat			
6.	Menentukan	6. text 7	<i>Text 7</i> is kind of	
	gambaran umum,	Heroine With Cigarette	recount text.	
	informasi tertentu,	Lighter	Recount text is	
	informasi rinci		Recount is a text	
	tersurat		which retells	
	torsurat	Question Number 37:	events or	

*****	What's the writer's	experiences in the	
	aim to write the above	past	
	text?	QN	
	Question Number 38:	37:menentukan	
	The underlined word	informasi rinci	
	has the same	tersurat.	6%
	meaning	QN	
	Question Number 39:	38:menentukan	
	Which is correct about	informasi rinci.	
	the text?	QN	
		39:menentukan	
		informasi	
7. Menentukan		tertentu.	
gambaran umum,			
informasi tertentu,			
tujuan			
komunikasi			
8. Menentukan	8. text 2	Text 2 is kind of	
gambaran umum,	THE NECTAR	descriptive	
makna kata,	The nectar is another	text.descriptive	
informasi tertentu,	useful secretion	text is a text to	

			T	
	informasi	Question Number 19:	describe person,	
	rinci	the extrafloral	thing	
		nectarines occur on the	QN	
		followings, except	19:menentukan	
		Question Number 20:	informasi rinci	
		what's the purpose of	tersurat.	
		the writer to write the	QN	
		text?	20:menentukan	8%
		Question Number 21:	gambaran umum.	370
		The synonym of the	QN	
		underlined word is	21:menentukan	
		Question Number 22:	makna teks	
		The following	tertulis.	
		statement are true,	QN	
		except	22:menentukan	
			informasi rinci	
			tersurat.	
		Text 6		
		ACACIA mimosaceae	Text 6 is kind of	
		The common name	descriptive text.	
		"watle by which these	descriptive text is	
,				

	plan are known in	a text to describe
	Australia came	person, thing
	into	QN
	Question Number 33:	33:menentukan
	The main purpose of	gambaran umum. 8%
	the writer to write the	QN
	text above is	34:menentukan
	Question Number 34:	makna teks
	The underlined word	tertulis.
	means Question Number 35:	QN
	The followings are the	35:menentukan
	characteristics of Phyllodes	informasi rinci
	except	tersurat.
	Question Number 36:	QN
	Which is not true about	36:menentukan
	the text?	informasi rinci
		tersurat.
9. Menentukan gambaran umum,	9. text 4 Advantages and	Text 4 is kind of hortatory
informasi tertentu,	Disadvantages of TV.	exposition.
tujuan	Question Number 26:	Hortatory
	What's the main idea	exposition is a

komunikasi	of paragraph 2?	text which gives	
Komumkasi		text which gives	
	Question Number 27:	opinion about a	
	The underlined word	problem by	
	means	persuading the	
	Question Number 28:	reader	
	The aim of the writer	QN	
	to write the above text	26:menentukan	
	is	gambaran umum.	
	Question Number 29:	QN	8%
	The underlined word	27:menentukan	
	means	makna kata dari	
		teks essai.	
		QN	
		28:menentukan	
		informasi rinci.	
		QN	
		29:menentukan	
		makna kata dari	
		teks esai.	
		Text 8 is kind of	
	Text 8		
Harry Potter order o	Harry Potter:order of	analytical	
	120.7 2 0001.01001 01	exposition.	

		the Poenix	Analytical	
		I absolutely love	exposition is a	
		Question number 40:	text which gives	
		What's the	oponion about the	
		communicative	problem withouth	
		purpose of the above	any	
		text?	persuading	
		Question Number 41:	QN	
		The underlined word	40:menentukan	
		means	gambaran umum.	6%
			QN 41:	0 70
		What does the	menentukan	
		underlined word refer	makna kata dari	
		to?	teks esai.	
			QN 42:	
			menentukan	
			makna kata dari	
			teks esai.	
·	10. Menentukan			
	gambaran umum,			
	informasi rinci			
	tersurat,			
	torburat,			
		·		

	T		
11. Menentukan	11. text 10	Text 10 is kind of	
gambaran umum,	In order to request	message. This	
rujukan kata dari	telephone service in	message used for	
teks esai	United States	who will instal a	
	Question Number 46:	telephone	
	The underlined word is	QN	
	closest meaning to?	46:menentukan	
	Question Number 47:	rujukan kata dari	
	When we order	teks essai.	4%
	telephone service who	QN	
	will install it?	47:menentukan	
		gambaran umum.	
	Text 11	Text 11 is not	
	Homework	appropriate with	
	I have been wondering	the indicators of	į
	if homework is	SKL since it does	
	necessary	not belong to a	
	Question Number 48:	specific text. It	
	The purpose of the text	can be	

P		
is	explanation text	
Question Number 49:	and can be	
What's the advantages	descriptive text.	4%
of doing the	QN	
homework?	48:menentukan	
	gambaran umum.	
	QN	
	49:menentukan	
	informasi tertentu	
	dari teks.	
Text 9	Text 9 is not	
The Shapered Boy	appropriate with	
There was once a	the indicators of	
shapered boy who	the SKL since it	
tended his sheep at	kind of anecdote	
the	text.	
Question Number 43:		
The underlined word	QN	6%
means	43:menentukan	
Question Number 44:	makna kata dari	
What lesson can we	teks.	
take from the story	ON	
Question Number 45:	QN	

Why did the boy like	44:menentukan	
calling out	informasi	
"Wolfwolf"?	tertentu dari	
	teks.	
	QN	
	45:menentukan	
	informasi rinci	
	dari teks.	
Text 3	Text 3 is not	
Free Time travel	appropriate with	
Agency	the indicators of	
Question Number 23:	SKL since it	
What the information does this chart shows	kind of	
us?	information	
Question Number 24:	text.	6%
Which vacation place	QN	
is the most popular?	23:menentukan	
Question Number 25:	gambaran	

	T	T	
	į į	How did the agency	итит.
	!		
	!	get this information?	ONI
			QN
			24:menentukan
			informasi rinci
	1		
			tersirat.
			QN
			25:menentukan
			informasi
			injormasi
(see			tertentu.
11 12		(see appendix 2)	
appendix 1)	(see appendix 1)	appendix 1) (see appendix 2)	

From the result above, in reading section, most of the items in the English try-out test for the final examination(UN) 2010/2011 published by National Education of Surabaya are appropriate with the SKL. For example from text 5 for number 30, 31, 32:

Text 5

This parking garage is closed for repairs. You can park 24 hours a day in the garage on Second Avenue between a.m and 9 p.m. or in the lot at the bus station across the street 24 hours a day. This garage will be reopen on the first of next month. We apologize for any inconvenience.

30.	whv	is	the	garage	is	closed?
50.	*****	13	uic	garage	13	CIUSCU:

- A. It is night time.
- B. It is full.
- C. It is being fixed.
- D. It is for buses only.
- E. It is bunkrupt.

31. What is sugeessted to the reader?

- A. Park in the street.
- B. Take the bus.
- C. Park in another garage or lot.
- D. Arrive earlier.
- E. Park outside the garage.

32. What is the type of above text?

- A. Announcement
- B. Advertisement.
- C. Memo.
- D. Article.
- E. Notice.
- This text is relevant with the SKL, is kind of announcement or notice of parking garage is closed for repair and relevant with the SKL. "Menentukan gambaran umum, informasi tertentu, informasi rinci tersurat, informasi rinci

tersirat, makna kata dari teks tertulis seperti teks fungsional pendek berbentuk announcement"

B. Discussion

From the result of the face validity, we can see that there are four criteria from its looks. First is how is the test's cover(this test's cover has blue color, good cover and interesting), second is how the size of letter used (the size of this test is 12 and can be read easily), third is how the test's layout (the layout of this test is tidy), and the last is how the size of paper used (this test use a legal size paper). From the directions. There are three criteria. First is how the general instruction of the test is (this test use general instruction and understanable), second is how the specific instruction every part is (every part instruction of this test is clear), and the third is how the instruction for going on to the next section or to the next page (there is no instruction going on next page or section, but in this test use the ending of section). From the items type shows that there are two criteria. The first is how the type of the test items choosen is (the type of this test multiple-choice test items). the second is how are the text presented (this text of the test is presented quite well). From the explanation above, we can get the result that English try-out test for the final examination(UN) 2010/2011 published by National Education of Surabaya is proper to conduct to the student.

From the result of the content validity, can see in listening section there are two kinds of test items, there are the item test are appropriate and not appropriate with the

SKL. In listening section, which not appropriate with the indicators in the SKL only number 5 and which the test items are appropriate with the indicators in the SKL are number 1,2,3,4,6,7,8,9,10,11,12,13,14,15. For reading section, which the test not appropriate with the SKL there are 3 texts. First, text 3 consist of number 23,24,25. Second, text 9 consistof number 43,44,45. And third text 11 consist of number 48,49.

According to Bloom, if the test agreement is 75% or more, then it can be said that the test has high content validity. On the other hand, if agreement is less than 50% the test is considered having low content validity.¹

After analyzing the content validity for instance the writer found the result that many test items in test items in the test are appropriate with the SKL. The test consists of fifty questions so we can say that every test item contains 2%.

From the explanation about the appropriateness of the test, English try-out test for the final examination(UN) 2010/2011 published by National Education of Surabaya has high content validity because from the whole questions (100%), 82% is appropriate with the SKL.

¹ Bejamin Bloom S, Handbook on Formative and Summative Evaluation of Student Learning, 1981, (NewYork: McGraw-Hill Book Co,), page 73.

CHAPTER V

CONCLUSIONS AND SUGGESTIONS

A. Conclusions

After analyzing the English try-out test items for UN 2010/2011 published by DIKNAS Surabaya in the previous chapter, it can be concluded that:

- The face validity of the English try-out test for UN 2010/2011 published by DIKNAS Surabaya is has a good face validity, from the test looks, from the direction which is used in the test, from the type of the test items shows that the test have good characteristic.
- The content validity of the English try-out test for UN 2010/2011 published by DIKNAS Surabaya has high content validity. Because from the test items are more than 50% appropriate with the SKL (Standard of Graduates' Competence).

B. Suggestions

There are some suggestions related to this result of the study below:

1. To the test constructor should pay attention with the test items which does not appropriate with the SKL. Eventhough the content validity of English try-out test items for UN 2010/2011 published by DIKNAS Surabaya have high

- content validity. The test contructor is suggested to revise it before it been tested.
- Test constractor and teachers should pay attention too to the instrument in conducting the test such as headset for listening section or tape recorder.
 Because this the important one for the student to listen well in conducting listening section.
- 3. To the teachers who carry out the test should monitor the students in doing the test and there is no deceitfulness in finishing the test. By doing the test right way, the try-out will proper to see students' readiness in facing the UN.