

**AN ANALYSIS OF ESSAY TEST ON ENGLISH FINAL
TEST FOR GRADE ELEVEN STUDENTS OF SMAN I,
LAMONGAN**

THESIS

**”Presented as Partial Fulfillment of the requirement for Attainment
the Degree of English Education”**



By:
ITA FARADILLAH
D05208116

**ENGLISH EDUCATION DEPARTMENT
FACULTY OF TARBIYAH
STATE INSTITUTE OF ISLAMIC STUDIES
SUNAN AMPEL
SURABAYA
2012**

APPROVAL SHEET

This thesis by ITA FARADILLAH entitled “AN ANALYSIS OF ESSAY TEST ON ENGLISH FINAL TEST FOR GRADE ELEVEN STUDENTS OF SMAN I, LAMONGAN” has been approved by the thesis advisors for further approved by the Board of Examiners.

Surabaya, September 14th 2012

Advisor I



Masdar Hilmy, MA. Ph.D
NIP 197103021996031002

Advisor II



Fitriah, MA
NIP 197610042009122001

EXAMINERS APPROVAL SHEET

Thesis Entitled:

“AN ANALYSIS OF ESSAY TEST ON ENGLISH FINAL TEST FOR GRADE ELEVEN STUDENTS OF SMAN I, LAMONGAN” has been accepted and approved by the broad examiners of English Department of Tarbiyah Faculty State Institute for Islamic Studies Sunan Ampel Surabaya

Surabaya, September 14th, 2012



Dean of Tarbiyah Faculty,

the
Dr. H. Nur Hamim, M.Ag

NIP. 196203121991031002

The Board Examiners
Advisor/Chair Person,

Masdar Hilmy

Masdar Hilmy, MA. Ph.D

NIP. 197103021996031002

Secretary,

Sigit Pramono Jati

Sigit Pramono Jati, M.Pd

NIP. 198605182011011005

Examiner I,

Mokhammad Syaifuddin

Mokhammad Syaifuddin, M.Ed

NIP. 197310131997031002

Examiner II,

M. Hanafi

M. Hanafi, MA

NIP. 197408042000031002

B.Heaton said it can be made as easy or as difficult as the test constructor wishes, it is frequently criticized simpler to answer than subjective test.

However, there are a lot of things need to be considered in creating essay test such as having a good validity of content, analyzing index difficulty and index discrimination. In creating a test, it is also crucial when it tests English language, test writers do not only write a test just it is but also try to conduct an applicable test.

There are several efforts of testing techniques that need to be defined before writing an acceptable test. A test can be said a good test and satisfactory if it contains the criteria of validity, reliability and the items of test are not too difficult and not too easy. As Hughes said that language ability is not easy to measure. There are some ideas to create good essay, a teacher should recognize some characteristics of good test, but it is not enough to measure whether an essay items are good or not, or when the test should be constructed. Another term that should be considered is the way of judging whether it is appropriate or not, it will be analyzed through index difficulty and index discrimination.

The writer chooses senior high school students since they are used to express their critical thinking or ideas in responding on essay test. Moreover the students are at grade eleven. The writer also thinks that SMAN I, Lamongan has international standard curriculum (RSBI), so the test of this school is made by the

test?. Therefore, face validity is almost perceived in the terms of content. If the test samples cover what learners have achieved or expect to achieve, the face validity will be perceived.

2. Reliability

Another way to measure the quality of good test is reliability. Reliability is consistent and dependable.⁵⁰ In short, to be reliable, a test must be consistent in its measurement. In other word, it measures the test or re-test reliability to distinguish from mark or re-mark reliability. For example: the teacher who gives the same test to the same subject or matched subjects on different occasions, then the extent that it produces similar results, it is reliable test. But if it is different results, means not reliable. This case as Haris stated on his book,

“Test reliability is affected by a number of factors, chief among them being adequacy of the sampling of the task”⁵¹

In line of that, Heaton said that it is primary importance in the use of both public achievement and proficiency test *and* classroom tests. For essay test, reliability is necessary to distinguish between three types of reliability: scorer, examinee, and examination. Each type of reliability emphasizes or accounts for certain types of errors.

⁵⁰ Douglas Brown, *Teaching by Principles : An Interactive Approach To Language Pedagogy*, (San Fransisco State University: Longman, Inc, 2001). 386

⁵¹ David P Harris, *TestingEnglish as a second Language*, (New York: Mc Craw-ill, Inc, 1959).

Table I: The example of analyzing Content Validity⁷

STANDARD COMPETENCE	BASIC COMPETENCE	INDICATORS	Learning Experience	ITEM TEST	Σ	%
MENDENGARKAN Memahami makna dalam percakapan transaksional dan interpersonal dalam kontek kehidupan sehari-hari	Merespon makna dalam percakapan transaksional (to get things done) dan interpersonal (bersosialisasi) resmi dan tak resmi secara akurat, lancer dan berterima yang menggunakan ragam bahasa lisan sederhana dalam berbagai kontek kehidupan sehari-hari dan melibatkan tindak tutur: berterima kasih, memuji, dan mengucapkan selamat	<ul style="list-style-type: none"> • Mengidentifikasi kata yang didengar, makna kata, hubungan antar pembicara. • Mengidentifikasi makna tindak tutur, berterima kasih, memuji, mengucapkan salam. • Merespon tindak tutur, berterima kasih, memuji, dan ucapan selamat 	<ul style="list-style-type: none"> • Mendengarkan berbagai tindak tutur yang didengar melalui tape atau teman • Mendiskusikan berbagai tindak tutur yang didengar melalui tape atau teman 	2, 5, 10, 7, 55	5	9%
				4, 6, 8, 11, 20, 21, 24, 28, 30, 34, 35	11	20%

⁷ Millatul Islamiyah, The content Validity and Item Analysis of semester II English Final Test for tenth grade students of SMAN 3 Sidoarjo, Unpublished S-1 thesis (Surabaya: IAIN Sunan Ampel-Tarbiyah 2010). Page 45

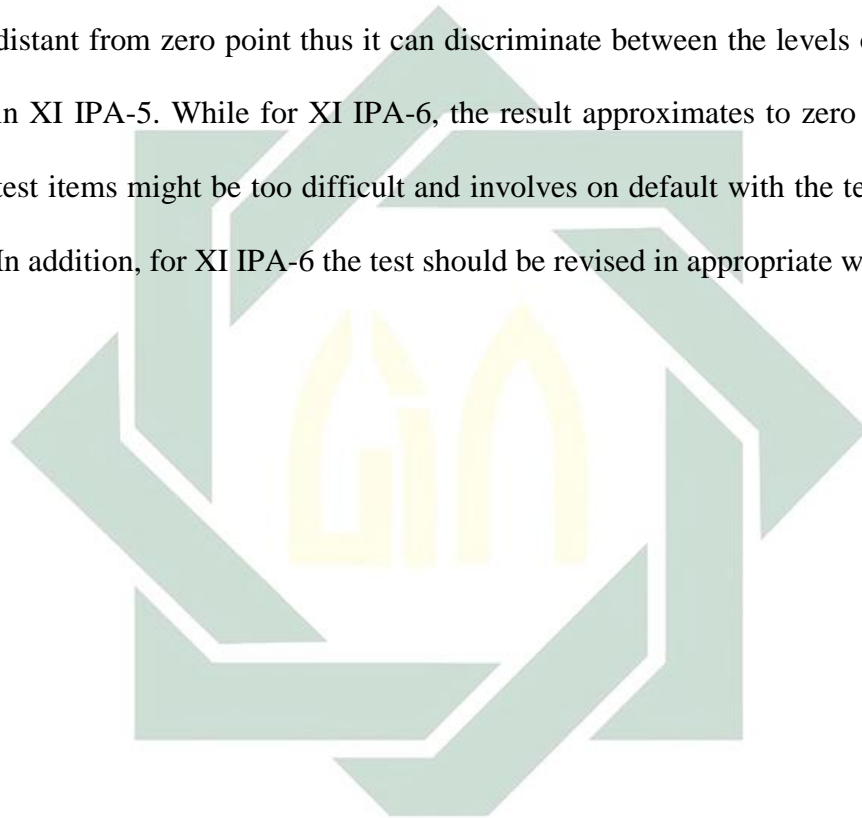
To compute index discrimination, first the writer compute index difficulty from each group, and then from the calculation of those groups the writer uses the formula of index discrimination. The result of index discrimination directly relates to index difficulty, if the value of index difficulty is good means that the test items work properly to discriminate the level of students' achievement. In this case, English essay test which has been evaluated for XI IPA-5 has acceptable discrimination. It means that test items used by SMAN I, Lamongan for XI IPA-5 work properly to differentiate between upper students and lower students, and can be functioned for other tests.

2. The Result of Analyzing Index Discrimination on Essay Test for XI IPA-6 class

In accordance with previous class, the calculation of index discrimination is primarily compute index difficulty between highest and lowest score of students' achievement on essay test. Then the result of higher group is diminished by the result of lower group as shown in appendix 9.

The table 26-30 in appendix 9 reports that index discrimination of test items used to evaluate English essay test for XI IPA-6 class is poor. Almost the test numbers are categorized as too easy and too difficult. Their value are around 00-0,20. There are 4 test numbers are considered as poor discrimination and 1 number could be said as satisfactory. It is 0,39% in number 35. The result indicates that the essay test items are not accepted to

Having discussed about index discrimination for students' of XI IPA-5 and XI IPA-6 in SMAN I, Lamongan , the writer concludes that test items made by the teacher in that school has acceptable discrimination for XI IPA-5 class because the test items are not too easy or too difficult. The result is distant from zero point thus it can discriminate between the levels of students in XI IPA-5. While for XI IPA-6, the result approximates to zero point. The test items might be too difficult and involves on default with the test making. In addition, for XI IPA-6 the test should be revised in appropriate way.



represent all materials taught in the class. Even though there was an item of test does not match with the content standard of English lesson for grade eleven students, it could say that the content of test items deal with several materials taught.

2. The result of index difficulty of Essay Test on English Final Test for Grade Eleven of SMAN I, Lamongan used by XI IPA-5 is acceptable. Almost the result in each test items are around 0,28-0,80, it means that the test item is good to evaluate students' achievement. Thus, those test items are safe and proper to be used for future test without being rewritten. While, the result of index difficulty for XI IPA-6 in each test items are around 0,1-1,00 which indicates that the test items are both too difficult and too easy. The test items should be revised for XI IPA-6 to evaluate the levels student's achievement. Too easy and too difficult items do not discriminate students' achievement well because when too easy items occurred, most of students get correct answer. Meanwhile, when the items are too difficult, then a chance to get correct answer is only for upper students.
3. The result of index discrimination for XI IPA-5 is acceptable. Between the result of index difficulty the result of index discrimination shows a balance. Most of test items are categorized to comply the criteria of high discrimination. It could say as good discrimination since the result is distant from zero point and close to point 1. In fact, the discrimination value is around 0,39-1,00. The different result occurs for XI IPA-6, most of the test

- Heaton, J. B. *Writing English Language Test*. New York: Longman Group. 1988.
- Islamiyah, Millatul. Content Validity and Item Analysis of Smester II English Final Test for Tenth Grade Students of SMAN 3, Sidoarjo. Unpublished S-I Thesis. Surabaya: IAIN Sunan-Ampel (English Education Department). 2010.
- Madsen, Harold S. *Techniques in Testing*. New York: Oxford University Press. 1983.
- Mardalis, Drs. *Metode Penelitian*. Jakarta: Bumi Aksara. 1995.
- Mehrens, William A. – Lehman, Irvin J. *Standardize Test in Education*. USA: Rinehart and Winston, Inc. 1969.
- Nasution, S. *Metode Research*. Jakarta: Bumi Aksara. 2006.
- Nurgiyantoro, Burhan. *Penilaian Dalam Bahasa dan Sastra*. Yogyakarta: BPFE Yogyakarta. 2001.
- Rahmatun, Ema Hafsah. An Analysis of English Multiple Choice Questions Test for Seventh Grade of SMP Buana Waru Sidoarjo. Unpublished S-I Thesis. Surabaya: IAIN Sunan-Ampel (English Education Department). 2011.
- Sudijono, Anas. *Pengantar Evaluasi Pendidikan*. Jakarta: PT. Raja Grafindo Persada. 1996.
- Thoha, Drs. M. Chabib. *Teknik Evaluasi Pendidikan*. Jakarta: PT Raja Grafindo Persada. 1991.
- Trochim, William, M.K. *The Research Method Knowledge Base*. Ithaca, N.Y: Cornell Custom Publishing. 1999.
- Tuckman, Bruce W. *Measuring Educational Outcomes, Fundamental of Testing*. USA: Harcourt Brace Jovanovich, Inc. 1975.