# CHAPTER I

# INTRODUCTION

This chapter discusses the area of the study that will be covered in some headings, background of the study, research questions of the study. Objectives of the study, significance, scope and limitation, then definition of key terms.

## A. Background of Study

These days, English has turned into the most frequent language used in more than 300 countries. This phenomenon results in the huge of need in English mastery in all fields including education[1]. For non-native English country for example, there are some English proficiency test for proving the English's ability of the test-takers such as TOEFL and IELTS.[2] Researcher believes that TOEFL and IELTS is urgently needed for English skill measurement.

The TOEFL is the most widely respected English-language test in the world.[3] It is administered to approximately 800 .000 candidates in more than 200 countries each year. More than 4.200 academic institutions government agencies, scholarship programs, and licensing/certification agencies in more than 80

---

[1] Aina, Qorry (2016) *AN ANALYSIS OF CONSTRUCT VALIDITY OF TOEFL-LIKE TEST IN ENGLISH INTENSIVE COURSE PROGRAM OF UIN SUNAN AMPEL SURABAYA.* Undergraduate thesis, UIN Sunan Ampel Surabaya. Page 1

[2] Brown, D. (2004). *Language Assessment Principles and Classroom Practices.* New York: Longman Press. Page: 72

[3] www.ets.org/toefl (The Official Web of TOEFL by ETS). Accessed on April 24th, 2016.

countries use TOEFL scores.[4] The Author of Educational Testing Service also adds that TOEFL is used in all over the world to test the English proficiency of people who live in non-English speaking countries. Because it is wide usage and internationally recognized test, TOEFL is used in all over the world include in Indonesia[5]. Indonesia is a country which its citizenry take English as third, or even foreign language. This is the reason why TOEFL is urgently needed.

In UIN Sunan Ampel Surabaya, TOEFL used is in form of equivalent test: made from the collection of test items from several resources, such as Cliff's TOEFL and Longman[6]. The TOEFL Equivalent Test is held by Language Development Center (P2B) of UIN Sunan Ampel Surabaya. There are two types of TOEFL Equivalent Test. The first is the test which used for the regular students who willing to take the post graduate and doctoral degree program in UIN Sunan Ampel Surabaya. The second is TOEFL Equivalent Test as the final examination in English Intensive Program. It is an appropriate policy of Language Development Center for standardizing students' English proficiency using TOEFL which is already used worldwide.

However, the reliability of TOEFL Equivalent Test by Language Development Center has never been examined before[7]. Regrettable, reliability is an obligation

---

[4] Ibid… Brown, Page: 84
[5] Rahmawati, Elis (2014) *AN ANALYSIS OF TEST-TAKING STRATEGIES USED IN TOEFLEQUIVALENT TEST BY SIXTH SEMESTER STUDENTS OF ENGLISH TEACHER EDUCATION DEPARTMENT UIN SUNAN AMPEL SURABAYA*. Undergraduate thesis, UIN Sunan Ampel Surabaya. Page 28
[6] Aina… Ibid. Page 1
[7] Aina… Ibid. Page 3

2

to ensure test development from time to time[8]. In addition, reliability also means test consistency. Without reliability, the consistency of a test is bias and undependable[9]. In the long term, the test would be untrustworthy and pointless[10]. Raising this issue, the researcher will conduct a study that examines reliability and the subject is TOEFL Equivalent Test by Language Development Center.

In language testing issue, reliability is the center of a test enterprise[11]. Reliability is a must that every single test should have especially if the test is the high-stakes one. High-stakes assessment situations are admission tests for universities or other professional programs, certification exams, or citizenship tests[12]. The placement test is qualified as one of the high-stakes assessment, for the test-takers of this test are the entire first year students.

In this study, the researcher will be focusing on analyzing the internal consistency reliability of the test at English Intensive Course Program which has never been examined[13]. Since reliability is one of a major issue in relabeling large-scale standardized test of proficiency[14]. Internal consistency reliability focuses on

---

[8] Lawrence, D. (2011). Reliability and Comparability of TOEFL iBT Scores. TOEFL iBT Research: Series 1, Volume 3. Page 3.

[9] Haertel, E. H. (2006). Reliability. WEsport, CT: American Council on Education and Praeger. Page 28.

[10] Zhang, Y. (2008). Repeater Analysis for TOEFL iBT. ETS Research Report (RM-08-05). Princeton, NJ: ETS. Page 85.

[11] Flucher, Glenn (2010). Practical Language Testing. United Kingdom: Hodder Education. Page 19.

[12] C. Roever. (2010). "Web-based Language Testing". Language Learning and Technology. Vol. 5 No. 2, Page 86.

[13] Aina… Ibid. Page 3

[14] Aina… Ibid, 25.

measuring reliability used to evaluate the degree to which different test items that probe the same construct produce similar result[15].

There are a wide variety of internal consistency measurement that can be used such as: 1) average Inter-item correlation, and 2) split-half reliability[16]. Average inter-item correlation is obtained by taking all of the items on a test that probe the same construct, determining the correlation coefficient for each pair of items, and finally taking the average of all of these correlation coefficients[17]. This method needs re-test which the researcher believes is inefficient because of the practical consideration. Inefficient practical consideration means the impossibility for P2B to give the second TOEFL Equivalent test just in order to measure the internal reliability. Therefore the split-half reliability is the most efficient way to test the internal reliability itself.

Split-half reliability is another subtype of internal consistency reliability. The process of obtaining split-half reliability is begun by ''splitting in half'' all items of a test that are intended to probe the same area of knowledge in order to form two ''sets'' of items. The entire test is administered to a group of individuals, the total score for each set is computed, and finally the split-half reliability is obtained by determining the correlation between the two total ''set'' scores[18]. Researcher

---

[15] Brown… Ibid page 124.
[16] Brown… Ibid Page 125.
[17] Cozby, C. (2001). *Measurement Concepts. Methods in Behavioral Research*. California: Mayfield Publlishing Company. Page 231
[18] James. D. B (2009). *What Is Internal Consistency Reliability?*" Shiken: JALT Testing & Evaluation SIG Newsletter. Vol 4: No: 2 Page 9.

4

believes that this way is the most effective and efficient way to get the data for internal consistency analysis.

Therefore, this research is conducted to analyze the internal consistency reliability of the Equivalent Test conducted Language Development Center UIN Sunan Ampel Surabaya.

## B. Research Problem

In relation to the background of the study previously outlined above, the problem of the study can be formulated as this following question:

What is the internal consistency reliability value of the TOEFL Equivalent Test conducted by Language Development Center of State Islamic University Sunan Ampel Surabaya using split-half method?

## C. Objective of The Study

The objective of the study of this research is to find the internal consistency reliability value using split-half method. International standard used for analyzing data is Spearman-brown Formula which measure internal consistency precisely.

## D. Significances of The Study

By conducting this research, the researcher hopes that it will give many benefits for the Language Development Center (P2B) of UIN Sunan Ampel Surabaya, English Intensive Program Lecturer, and further researcher.

1. For Language Development Center (P2B) of UIN Sunan Ampel Surabaya

5

By conducting this study, the researcher hopes it can help Language Development Center (P2B) of UIN Sunan Ampel Surabaya in standardizing the TOEFL at Intensive English Program for the internal consistency of the test has never been investigated before. Moreover, the Language Development Center (P2B) of UIN Sunan Ampel Surabaya may also use the result of this research as the basic consideration in constructing the test for the next following years.

2. For the English Intensive Program Lecturer

The researcher hopes this study will be one of the consideration for creating English Intensive Program material. The consistent TOEFL Equivalent test item produces high quality questions. They can be added in to English Intensive handbook.

3. For Further Researcher

This research can be used as the basic reference in conducting another analysis of internal consistency reliability that deals with a test, especially English Language Proficiency Test.

E. Scope and Limits

1. Scope of the Study

In language testing, there are several kinds of reliability. However, the researcher confines this research to examine the internal consistency reliability of the TOEFL at English Intensive Class Program in UIN Sunan Ampel Surabaya. This will not observer neither the students' performance on

6

the class nor the class administration. This research will focus on finding the internal consistency reliability only.

2.  Limits of the Study

    This research is limited to investigate the TOEFL which is used in English Intensive Class Program as final examination.

## F.  Definition of The Key Terms

1.  Internal Consistency Reliability

    Internal consistency reliability is an assessment of how reliably survey or test items are designed to measure the same construct. In specific, a construct is an underlying theme, characteristic, or skill such as reading comprehension or customer satisfaction[19]. There are a variety of internal consistency measures[20]. They are average inter-item, and split-half method. The method used in this research is split-half.

2.  Split-Half Method

    As this research focuses on examining the internal consistency reliability value, the researcher uses split-half method. Split-half method is a way of dividing the test items or survey into two different score-items. Therefore, the coefficient correlation value can be observed between the gaps.

---

[19] Brown, D. (2004). *Language Assessment Principles and Classroom Practices*. New York: Longman Press. Page: 124

[20] James. D. B (2009). *What Is Internal Consistency Reliability?* Shiken: JALT Testing & Evaluation SIG Newsletter. Vol 4: No: 2 Page 9.

3. TOEFL Equivalent Test at English Intensive Program

TOEFL at English Intensive Program is a test made by P2B of UIN Sunan Ampel Surabaya which use TOEFL as the standard in giving scores and making the questions. P2B does not make the test items by themselves, they take the questions from various references such as Cliff's TOEFL. This test is divided into three sections: listening, grammar, and reading. The tests-takers of this test are the first year students of UIN Sunan Ampel Surabaya. The minimum score of this test is 400. If they fail in reaching the minimum score, they can re-take the test until they get the minimum standard. The certificate of TOEFL by at English Intensive Program is used as one of the requirements for participating thesis examination.

3. Intensive English Program

Intensive English Program is a pre-academic program which is designed to prepare the students for a regular English course and improving the students' English competence[21]. The English Intensive Program defined in this study is a must-take class for the first year students of UIN Sunan Ampel Surabaya.

---

[21] H. Douglas Brown. (2000). *Teaching by Principal: An Interactive Approach to Language Pedagogy*. California: Longman. Page 123

8