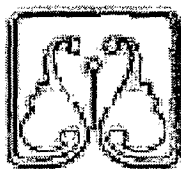


In the Name of God
The Merciful
The Compassionate

131/14



Guilan University
Faculty of Literature and Humanities
English Language Department

**A Thesis Submitted in Partial Fulfillment of Requirement
for the Degree of Master of Art in Teaching English as a
Foreign Language (TEFL)**

Title:

*The Effects of Alternative Assessment Methods on
Proficiency Level of High school Students*

By:

S. Hamidreza Aliaki

Supervisor: Dr. Hooshang Khoshsima

Advisor: Dr. Behzad Barekat

September, 2009



۱۴۱۷۱۶

We hereby approve that this thesis by Seyed Hamidreza Aliaki entitled “The Effects of Alternative Assessment Methods on Proficiency Level of High school Students” be accepted in fulfillment of the requirement for the degree of Master of Arts in Teaching English as a Foreign Language.

Sponsoring committee

Dr. Hooshang Khoshsimia (Thesis Supervisor)

.....

Dr. Behzad Barekat (Thesis Advisor)

.....

Dr. Masood Khalili (Head of English Department)

.....

September 2009

Dedicated to

***My dear wife whose angelic being has
made my life full of beauty and confidence***

&

***My parents whose prayers always bring
success to me***

TABLE OF CONTENTS.....	PAGE
ACKNOWLEDGMENTS	VII
ABSTRACT.....	VX
CHAPTER 1	1
1-1 INTRODUCTION.....	2
1-2 STATEMENT OF THE PROBLEM.....	3
1-3 PURPOSE OF THE STUDY.....	5
1.4. RESEARCH QUESTION AND HYPOTHESIS.....	5
1-5 LIMITATION OF THE STUDY.....	5
1-6 OUTLINE OF THE STUDY.....	6
1-7 DEFINITION OF SOME TERMS.....	6
Assessment.....	6
Test.....	6
Formative Assessment.....	6
Summative Assessment.....	6
Alternative assessment:.....	6
Portfolio assessment:.....	7
Self assessment:.....	7
Feedback.....	7
Washback.....	7
CHAPTER 2	8
2-1 ASSESSMENT.....	9
2-2 TRADITIONAL VERSUS ALTERNATIVE ASSESSMENT.....	10
2-2-1 Traditional assessment.....	12
2-2-2 Alternative assessment.....	15
2-2-2-1 Definition.....	15
2-2-2-2 Some further terminology.....	16
2-2-2-3. Benefits of alternative assessment.....	17
2-2-2-4 Alternative Methods of Assessment.....	20
2-2-2-5 Concerns raised about certain qualities of alternative assessment.....	21
2-2-2-6 Responses to concerns raised.....	23
2-3 THE CONCEPT AND CHARACTERISTICS OF PORTFOLIO.....	24
2-3-1 Definition of portfolio.....	24
2_3_2 characteristics of portfolio.....	27
2_3_2_1 Physical Characteristics of Second/Foreign Language Portfolios.....	28
2_3_3 Advantages of portfolio.....	29
2_3_4 Disadvantages of portfolio.....	35
2_3_5 Implementation of portfolio.....	37
2_3_5_1 Portfolio assessment is useful for:.....	38
2_3_5_2 Portfolio assessment is not as useful for:.....	38
2_3_5_3 How to use portfolio assessment.....	39
2_3_5_4 Preparing Your Portfolio.....	39
2_3_6 portfolio: some description.....	41
2-3-6-1 Non selective portfolio.....	41
2-3-6-2 The selective portfolio.....	41
2-4 SELF ASSESSMENT.....	41
2-5 PEER ASSESSMENT:.....	43
2-5-1 Peer assessment in EFL contexts.....	46

2-6 ASSESSMENT, LEARNING AND FEEDBACK.....	48
2-7 RELIABILITY AND VALIDITY OF PORTFOLIO.....	50
CHAPTER 3	52
3-1 INTRODUCTION.....	53
3-2 DESIGN OF THE STUDY:.....	53
3-3 PARTICIPANTS	53
3-4 INSTRUMENTS:	53
3.5 PROCEDURES	54
CHAPTER 4	58
DATA ANALYSIS AND DISCUSSION	59
CHAPTER FIVE	65
5-1 CONCLUSION.....	66
5-2 PEDAGOGICAL IMPLICATIONS AND APPLICATIONS	66
5-3 SUGGESTIONS FOR FURTHER RESEARCH.....	68
REFERENCES	69
APPENDIXES	77

TABLE OF FIGURES AND TABLES.....	PAGE
TABLE 3-1 TIME TABLE FOR PRE TEST AND POST TEST	53
TABLE 3-2 GUIDELINES FOR SELF AND PEER ASSESSMENT.....	55
TABLE 3-3	56
TABLE 4-1	59
TABLE 4-2	60
TABLE 4-3	60
FIGURE 4-1.....	61
TABLE 4-4.....	62
TABLE 4-5.....	62
FIGURE 4-2.....	63
TABLE 4-6.....	64

Acknowledgments

I am really grateful to Dr. Hooshang Khoshshima, the supervisor of this thesis, for all his kind and support, guidance, encouragement and assistance during completing this research project.

I am also truly grateful to Dr. Behzad Barekat, the advisor of this thesis, for his much careful reading of this reading report. His insightful comments were of great help and value to the completion of this thesis.

My special thanks go to Dr. Khalili, the head of department, who provided me the best conditions for completing my thesis.

Last but not the least, I would like to express my warmest appreciation and thanks to my wife to whom I owe my life and courage.

Abstract

Title: The Effects of Alternative Assessment Methods on Proficiency Level of High school Students

Author: Seyed Hamidreza Aliaki

Assessment procedures have been changed in line with shifts in instructional paradigms. As language teaching and learning has moved in the direction of learner centered skills, assessment has begun to incorporate a wider variety of measures that more closely reflect the types of tasks students perform in classroom or their homes. So assessment procedures are becoming more authentic.

In order to monitor the progress of students, assessment needs to be conducted on an ongoing basis with procedures which effectively give us useful information on class instruction.

The present study has tried to investigate the wash back effect of alternative assessment techniques such as portfolio, self-assessment and peer assessment on high school students' English proficiency. So after examining traditional and new trend of assessment we tried to explore the effects of portfolio assessment, peer assessment and self assessment on the students' progress. These techniques were utilized to monitor the improvement of students' English proficiency.

A number of 52 students of high school in first grade from Imam Khomeini high school in Bakarz (a city in Razavi province) participated in this study. 26 students were served as experimental group and 26 others as control group; that is the participants were assigned into two groups. They attended two-hour-a-week English classes.

A set of t-tests were utilized to test the research hypothesis. The t-observed for post test was greater than t-critical. Based on the findings obtained in this research the effectiveness of the treatments was confirmed. It can be implied that that employing alternative assessment techniques in the assessment of students proficiency has led to a significant difference between the performance of the experimental and control groups

The results of this study may imply that developing the habits of self-assessment, critical and reflective thinking and providing the learners with the habit of feedback are effective techniques to be incorporated into the language classes

Key words: Assessment; alternative assessment; portfolio assessment; peer assessment; self assessment

Chapter 1

Introduction

Chapter 1

1-1 Introduction

Assessment is generally seen as one of the key challenges in the field of learning (Gardner & Miller, 1999; Richards & Rodgers, 2001; Brown, 2000). Many researchers and practitioners point to difficulties with assessing language gains in an environment in which variables cannot comprehensively be controlled for. Furthermore, assessing the development of learning gains, such as the ability to plan and manage independent learning or to approach a learning task strategically relies on underdeveloped methodologies and assessment tools.

One of the most challenging tasks for language instructors is finding effective ways to determine what and how much their students are actually learning. Instructors need to think carefully about what kinds of knowledge their tests allow students to demonstrate.

"Assessment and testing are important dimensions of every teaching and learning practice. They provide teachers with the necessary information about their students' language ability and enable them to make sound professional judgments and decisions within the context of their classes" (Bachman, 1995).

Language testing, generally associated with formal assessment procedures such as tests and examinations carried out at specified times and serving a variety of purposes (i.e. diagnostic, achievement, progress, etc.), is a vital component of instructional language programs throughout the world.

While this type of assessment is a mainstay of educational programs (Butterfield et al., 1999), educators and critics from various backgrounds have raised a number of concerns about its usefulness as the primary measure of student achievements.

In the past according to Tannenbaum (1996; cited in Coombe and Barlow, 2004) educators came to realize that alternative form of assessment are an important means of gaining dynamic picture of student's academic and linguistic development. One of the most exiting developments in the reform of teacher education programs is the use of alternative form assessment to evaluate student's learning and one of the most popular forms of alternative assessment is the use of portfolio (Barret 2001).

Portfolio assessment is the forefront of alternative assessment approaches (Coombe & Barlow 2004). According to Wiggins (1989 cited in Coombe & Barlow 2004) there has been a growing recognition that a single measure was incapable of estimating the diversity of skills, knowledge processes and strategies that combine to determine students' progress. Altan (2002) indicated that the most pervasive inadequacy in traditional and standardized

testing was the assumption that all students could be assessed by using the same instrument.

“Since we don’t learn in the same way we cannot be assessed in a uniform fashion. Therefore teachers must seek to assess their students’ learning in way which will give an accurate overview of their strength and weaknesses. (Altan, 2002)

Gardner(1987) expresses this view: “I believe that we should get away altogether from tests and correlations among tests, and look instead at more naturalistic sources of information about how peoples around the world develop skills important to their way of life”. He also suggests that "the school seeks to encourage the unique blend of intelligences in each of its students, assessing their development regularly in intelligence-fair ways”

Twentieth century gestalt psychologists like Kohler (1997) stated that “learning takes place through an act of insight while learner engages and reflects on experiences.”

There are two reasons for undesirability of traditional assessment as we see in our own education system in Iran. One is educational aspect and the other is psychological aspect. By the educational aspect, it is meant that our one-shot single final exams naturally make students rely on their memorization ability and merely try to memorize some pieces of information as the so-called content of the syllabus and reproduce these pieces of information from their memory on the exams to score high and after the exam this information disappeared. This is a kind of rote learning which is not meaningful at all. By the psychological aspect, it is meant that it is stressful for them when students know, from the beginning of the course, that only one final exam makes their fates and if they fail they will be back one year. Their attendance and activities during the term would be useless. This is, to great extent, because they already know that they will have a single opportunity to fight for their future once and for all. Whatever they have done so far and to whatever degree they may have been qualified will be of no help to them if they do not do well enough on a single exam either because they were sick the night before the exam or they were simply embarrassed by the examination situation. This is why they develop certain negative affective and psychological filters towards the teaching-learning-testing processes.

1-2 Statement of the problem

It is usual that teachers make end to their education in the semester by testing students at the end of each term. And everyone who is successful in this final exam will pass the course. This one shot examination which is very common in Iran does not give us enough

information about student's performance in real life and doesn't tell us how much he uses the information he has learned in real situations.

Traditional forms of assessment have often been criticized for their failure to truly assess children's learning. Traditional forms of assessment do not represent activities children typically perform in classrooms, they do not reflect current theories of learning and cognition, and they are not based on abilities students actually need for their future success (Pierce & O'Malley, 1992). In addition, traditional forms of assessment often focus more on a product versus the processes used in learning and creating. For example, objective paper-and pencil tests usually focus on whether children get the "right" answers, instead of the process of how they arrive at their responses (Fischer & King, 1995). Another concern about the traditional forms of assessment, such as tests, is that they cannot be used to closely monitor children's progress in the school curriculum throughout the year because they are only administered once or twice annually (Pierce & O'Malley, 1992). As a result, they do very little to guide teachers' planning and children's learning.

To ascertain a more realistic assessment of their children's abilities, many teachers are turning to forms of alternative or authentic assessment. Authentic assessment involves gathering information concerning a child's performance while the child is engaged in genuine or realistic learning opportunities (Rhodes & Shanklin, 1993). Various forms of authentic assessment have been used to supplement traditional testing. Authentic assessments have proven to be particularly useful in many classroom situations where paper-and-pencil tests cannot gather specific types of information about children's achievement (Tierney, Carter, & Desai, 1991). Some examples include not only assessments in the academic areas, such as reading and writing, but also in the areas of art and the development of social skills (Levick, 1998). To gather appropriate evidence related to such performances, teachers have to observe and judge each pupil's actual performance and the products of those performances (Airasian, 1994).

Performance assessment and portfolio assessment are two examples of alternative or authentic assessment strategies. These forms of assessment enable teachers to continuously assess children's progress with regard to the learning processes they use as well as the products they produce. These forms of assessment also allow children to use their higher-order thinking skills, and create a collaborative approach to assessment which enables teachers and children to interact in the teaching/learning process (Pierce & O'Malley, 1992).

1-3 Purpose of the study

The purpose of this study is to investigate the effect of alternative assessment techniques like portfolio assessment, peer assessment and self assessment on achievement learning English language of first grade students of high school.

Students are asked to utilize peer assessment, self assessment, portfolios and bring their tasks given to them. The students reflected upon their own learning processes. The self assessment portfolios were designed to enable students to become aware of both successful and unsuccessful strategies that they employed in learning English as foreign language.

The educational system in Iran has a traditional of rote learning and traditional assessment. These traditional testing and assessments are not authentic and don't demonstrate actual level of proficiency.

1.4. Research Question and Hypothesis

To investigate the effectiveness of alternative assessment techniques for improving language proficiency of Iranian high school student, the following questions were posed:

- 1- Is it possible to enhance students' achievement utilizing alternative assessment techniques?
- 2- In which skill alternative assessment techniques work better?
- 3- Is there any relationship between alternative assessment and students' motivation?
- 4- 1. Is there any relationship between alternative assessment techniques and language proficiency improvement of the Iranian high school students?

Accordingly the following alternative hypothesis was formulated:

There is a positive relationship between utilizing alternative assessment techniques and language proficiency improvement of the Iranian high school students.

The following null hypothesis was formulated:

There is no relationship between utilizing alternative assessment techniques and language proficiency improvement of the Iranian high school students.

1-5 Limitation of study

In this study we were in shortage of time and the school which we were teaching in was resistance to any change in number of class and neglecting final exam for the experimental group. We also had to select only from male students. The first grade students were selected for they have more sessions a week but their English proficiency was not so high and we had to translate some parts of peer and self assessment for them.

We also had to select the texts from their own books because the institute didn't let us to use our own.

1-6 Outline of study

This study consists of five chapters. Chapter 1 "introduction" states the problem which the researcher wants to investigate; purpose of the study; limitation of the study; and definition of some terms. Chapter 2 reviews some related literature and related researches that have already been done. In chapter 3 the methodology of this study is discussed which includes design of study; participants; instruments and procedures. Chapter 4 analyzes data gathered from this study and discusses the results. Chapter 5 is the conclusion of this research and it gives pedagogical implementations and some suggestions for further research.

1-7 Definition of some terms

Assessment:

Assessment is the process of documenting, usually in measurable terms, knowledge, skills, attitudes and beliefs. Assessment can focus on the individual learner, the learning community (class, workshop, or other organized group of learners), the institution, or the educational system as a whole

Test

A test is a measurement instrument designed to elicit a specific sample of an individual's performance (Bachman, 1995, p.20).

Formative Assessment

Formative assessment takes place during the course of teaching and is used essentially to provide feedback into the teaching / learning process (Gipps, 1994, p.vii).

Summative Assessment

Summative assessment takes place at the end of a term or a course and is used to provide information about how much students have learned and how well a course has worked (Gipps, 1994, p.vii).

Alternative assessment:

An assessment that requires students to generate a response to a question rather than choose from a set of responses provided to them. Exhibitions, investigations, demonstrations, written or oral responses, journals, and portfolios are examples of the assessment alternatives we think of when we use the term "alternative assessment." Ideally, alternative assessment requires students to actively accomplish complex and significant

tasks, while bringing to bear prior knowledge, recent learning, and relevant skills to solve realistic or authentic problems. Alternative assessments are usually one of the key elements of an assessment system.

Portfolio assessment:

Portfolio is a collection of work, usually drawn from students' classroom work. A portfolio becomes a portfolio assessment when (1) the assessment purpose is defined; (2) criteria or methods are made clear for determining what is put into the portfolio, by whom, and when; and (3) criteria for assessing either the collection or individual pieces of work are identified and used to make judgments about performance. Portfolios can be designed to assess student progress, effort, and/or achievement, and encourage students to reflect on their learning.

Self assessment:

Self-assessment *in an educational setting* involves students making judgments about their own work. Assessment decisions can be made by students on their own essays, reports, projects, presentations, performances, dissertations, and even exam scripts. Self-assessment can be extremely valuable in helping students to critique their own work, and form judgments about its strengths and weaknesses. For obvious reasons, self-assessment is more usually used as part of a formative assessment process, rather than a summative one, where it requires certification by others.

Feedback

Comments or information that learners receive on the success of a learning task, either from the teacher or from other learners (Richards, 1992, p.137).

Washback

Washback or backwash, also known as measurement-driven instruction, is a common term in applied linguistics referring to the influence of testing on teaching and learning, which is a prevailing phenomena in education..

Chapter 2

Review of Literature

Chapter 2 Review of Literature

2-1 Assessment

The root of the term "assessment" is *assidere*, which is also the root of the French *asseoir*, to seat or set. It was first used in the sense of setting the value of property to apportion a tax. Assessors traditionally make a site visit -- they inspect the property or the situation and its documents, they categorize its functions, they hear from the owner of the property, they evaluate it by setting it against already-existing standards, and so forth. The assessment requires time, as well as interaction between the assessor and the person or property being assessed, so that the congruence of perception with reality or, in our case, the congruence between underlying mental processes and surface observation, can be verified. The idea here is that the product is not sufficient evidence of the quality of the thinking processes that produced it.

Well designed assessment procedures are essential to meeting the needs of language minority students acquiring English as their second language. Assessment is involved at many steps in a continuum of services for these students: in initial identification' in placement of students into appropriate instructional programs, in monitoring the progress students make within these programs, in reassigning students to different levels within a program depending on their growth in English language skills, in moving students out of special programs and into mainstream classes, and in following the progress of these students in mainstream. This continuum is wholly dependent at each step on the appropriate selection, use, and interpretation of relatively complex assessment procedures. Students' learning is to a large extent affected by assessment tasks (Biggs, 1999). If the assessment task is perceived as a reproduction of previously learned content, students are more likely to adopt a surface approach to learning. On the other hand, if students perceive the assessment task as achieving personal understandings of the content, a deep approach to learning is more likely to be adopted. Empirical evidence exists to support the suggestion that different formats of assessment produce different approaches to learning (Scouller, 1998; Tang, 1994).

Assessment is perhaps one of the least understood areas of language teaching and learning, especially in Iran. Students and teachers alike seem to have problems with the word "testing". It seems that students see tests as a threat to their competence because they are afraid that they will not perform well. Teachers often do not like to construct tests and are not altogether satisfied with the results when they do. They are also suspicious of the standardized, professionally designed tests because they are not always sure what these

tests are actually trying to measure (Celce-Murcia, 2001). Some students and teachers alike argue that tests can not be indicative of one's language ability mainly because they are more product-oriented, and students' performances are measured and rated only based on the product. As a result, students usually get concerned of the exam result because they must be physically and psychologically all right to take their exam well. So, it would not be fair to those students who were in the process of learning during the semester but rather unwell on the time of the exam due to some factors, e.g. physical factors, environmental situation, etc.

As surveyed by Shohami (1985), the word "test" has been interpreted by students as follows:

- a. Tests are used as punishment.
- b. Tests are administered instead of teachers' giving instruction.
- c. The tests are the only measure for grading.
- d. Tests don't reflect what was taught.
- e. The tests are returned with a lack of corrections or explanation.

As seen in the students' responses above, both students and teachers are suspicious of tests. Fortunately, in recent years, there has been a growing interest in improving the situation in most countries. In Iran also, there has recently been a move toward some studies into the effect of assessment on the language teaching, though there is a long way to go.

As a form of assessment, portfolios are designed to encourage students to adopt a deep approach to learning. Despite the promises, there is relatively little research on the effect of portfolio assessment on students' approaches to learning. Therefore, one of the objectives for undertaking this study is to examine the effect of portfolio assessment as a branch of alternative assessment on learners' proficiency.

2-2 Traditional versus alternative assessment

One useful way to think about assessment is to contrast it with testing, an ever-present factor that confronts teachers and students in all disciplines. Tests have come to be an accepted component of instructional programs throughout the world. Sometimes tests are justified on the basis of accountability: are students learning what they are supposed to be learning? Decision-makers need this type of evidence in order to make judgments about how to spend resources, for example. Sometimes, tests are viewed as feedback for language students concerning their progress. Oller (1979) stated that "the purpose of tests is to measure variance in performances of various sorts." In this sense, testing--typically

achievement testing--serves as a monitoring device for learning. Tests are given at a particular point in time to "sample" student learning. Most of us are familiar with "paper and pencil" tests even if they take on a computerized format. Ordinarily, after the test is given, some type of reporting takes place, often in the form of a single score or grade. Sometimes, decisions are made based on test results (e.g., retake the test, pass the course, go on to the next unit of instruction, etc.). A final important aspect of testing is that the test is usually kept hidden from the students until it is administered, indicating a degree of secrecy in order to assure confidentiality.

Assessment then can be shown to be very different. Some important differences between testing and assessment become obvious. In an instructional program, assessment is usually an ongoing strategy through which student learning is not only monitored--a trait shared with testing--but by which students are involved in making decisions about the degree to which their performance matches their ability. Spolsky (1992, cited in Smith, 1999) argues that diagnostic or formative assessment is typically curriculum-driven. This type of assessment shadows the curriculum and provides feedback to student and teachers. She wisely argues, too, for a multilevel system that combines testing and assessment

A paraphrase of this model would go something like this:

- Students are provided opportunities before and after units of instruction to assess their own performance (self-assessment).
- Teachers periodically assess students' performance and both discuss their respective assessments (tests and measurements).
- Occasionally, some external monitor assesses the student's (and perhaps the teacher's) performance and discusses it with the teacher.

Assessment, then, should be viewed as an interactive process that engages both teacher and student in monitoring the student's performance. Criterion-referenced testing is clearly based on this way of relating teaching-testing-assessment for congruence. Language testing, generally associated with formal assessment procedures such as tests and examinations carried out at specified times and serving a variety of purposes (i.e. diagnostic, achievement, progress, etc.), is a vital component of instructional language programmes throughout the world. While this type of assessment is a mainstay of educational programmes (Butterfield et al., 1999), educators and critics from various backgrounds have raised a number of concerns about its usefulness as the primary measure of student achievements. Before attempting to discuss 'alternative assessment' at any