

## The Effect of an Instructional Program based on Authentic Assessment on the Achievement of EFL Secondary Stage Jordanian Students Reading

Mohammed Mohaidat \*

Received Date: 6/5/2012

Accepted Date: 17/12/2012

**Abstract:** The present study aimed at investigating the effect of an instructional program based on authentic assessment on the achievement of EFL secondary stage students in Jordan in the reading Skills through answering the following question:

- Is there a significant difference in Jordanian Secondary school students' achievement in reading comprehension between the experimental group and the control group due to the method of assessment (Traditional or Authentic)?

The study hypothesis was stated as follows:

- There is no statistically significant difference ( $\alpha = 0.05$ ) in Jordanian secondary school students' achievement in reading comprehension between the experimental group and the control group due to the method of assessment.

The subjects of the study consisted of 70 eleventh grade students who were divided into an experimental group and a control group.

To examine the study hypothesis, the researcher developed an instructional program based on authentic assessment techniques to assess the experimental group students. The researcher also developed another program based on traditional assessment to assess the control group students. In addition, the researcher developed an achievement test to assess the subjects' reading skills before and after conducting the experiment. Validity and reliability of the instruments were ensured. The experiment lasted for one semester (approximately three months). MANOVA and the t-test were used.

The findings of the study showed that there was a statistically significant difference between the two groups on the reading achievement post-test in favor of the experimental group which was taught using authentic assessment techniques.

In light of the findings of the study, the researcher proposed some recommendations to the Ministry of Education, EFL teachers, and researchers who are interested in authentic assessment.

Assessment of students' learning and progress has always been a challenge for educators. In the 20<sup>th</sup> century, traditional assessment was the most used type of assessment for measuring students' progress and achievement (Meadows and Karr, 2001). However, this type of assessment had its own limitations and weaknesses. This gave rise to calls for new ways to assess and report students' abilities and progress when students move at different rates toward individualized goals instead of marching toward grade-level standards as in the past

### أثر برنامج تدريسي معتمد على التقويم الواقعي على تحصيل طلبة المرحلة الثانوية في الأردن في مهارة القراءة في اللغة الانجليزية

محمد مهيدات: كلية الآداب، قسم اللغة الانجليزية، جامعة الملك سعود- الرياض

ملخص: هدفت هذه الدراسة الى استقصاء أثر برنامج تدريسي معتمد على التقويم الواقعي في تحصيل طلبة المرحلة الثانوية في مهارة القراءة في اللغة الانجليزية في الأردن.

وقد حاولت الدراسة الاجابة عن السؤال التالي:

- هل يوجد فرق في تحصيل طلبة المرحلة الثانوية في الأردن في مهارة القراءة نتيجة لطريقة التقويم؟

ولقد صيغت فرضية الدراسة كالآتي:

- لا يوجد فرق ذو دلالة إحصائية  $\alpha = 0.05$  في تحصيل طلبة المرحلة الثانوية في الأردن في مهارة القراءة تعزى لطريقة التقويم.
- وقد اختيرت عينة الدراسة من طلاب الصف الحادي عشر في مدرسة سعد بن أبي وقاص في مدينة اربد في الأردن. و بلغ عددهم 70 طالبا تم تقسيمهم الى مجموعتين: مجموعة تجريبية ومجموعة ضابطة.

ولاختبار الفرضية السابقة، فقد طوّر الباحث برنامجاً تدريسياً قائماً على التقويم الواقعي، وقد استخدمه لتقويم طلبة المجموعة التجريبية. كما طوّر الباحث برنامجاً اخر قائماً على التقويم التقليدي لتقويم طلبة المجموعة الضابطة. وبالإضافة إلى ذلك فقد طوّر الباحث اختباراً تحصيلياً لكلا المجموعتين لتقويم مهارتي القراءة والكتابة قبل وبعد إجراء التجربة. و تم تطبيق إجراءات التحقق من صدق البرنامجين التدريسيين كما تم التحقق من صدق وثبات الاختبار التحصيلي. دامت التجربة فصلاً دراسياً كاملاً (حوالي ثلاثة شهور). وقد تم تحليل نتائج الطلبة في إختبار القراءة باستخدام اسلوبي تحليل التباين المتعدد واختبار (ت). أظهرت نتائج الاختبار البعدي وجود فرق ذي دلالة إحصائية بين المجموعة التجريبية والمجموعة الضابطة في مهارة القراءة لصالح المجموعة التجريبية التي تم تدريسها باستخدام برنامج التقويم الواقعي.

وفي ضوء هذه النتائج التي توصلت إليها الدراسة قدم الباحث توصيات لوزارة التربية والتعليم ومدرسي اللغة الإنجليزية والباحثين المهتمين بالتقويم الواقعي.

There are many differences between authentic assessment and traditional assessment. The first and most important difference is each type's vision of learning. Traditional assessment assumes that learning involves memorizing discrete pieces of objective information in a linear, step-by-step fashion (Kohn, 1999; Rossi, 2000); that learning can be effectively assessed by demanding that desired pieces be produced

\* king saud university,college of arts, english department, Al- Riyadh, Kingdom of Saudi Arabia.

© 2013 by Yarmouk University, Irbid, Jordan.

in a timed, paper-and-pencil tests; that academic achievement can be measured and quantified by counting correct responses; and that nonacademic areas (such as the ability to work cooperatively with others, to make independent decisions, to effectively self-evaluate) where progress cannot be quantified are not worth attempting to assess (Caustad, 1996). These assumptions were derived, at least in part, from the behaviorist theories of learning that dominated psychology in the early part of the 20<sup>th</sup> century.

Derived from the constructivist theory of learning, and in contrast with traditional assessment, authentic assessment considers learning as a multi-dimensional process that is subject to great individual variation (Eisner, 1999). It is a process where children "construct" knowledge and modify previous understanding as they interact actively with their environment and other people, rather than absorbing knowledge passively; where individuals vary greatly the pattern of cognitive development and rely on different learning styles, and where learners' emotional state affects learning to a great degree (Burke, 1999).

Assessment of student learning has to be changed due to the fact that there are changes in the skills and knowledge needed for success in the world today. The learning goals for the students are changing as the current knowledge about the relationship between assessment and instruction has changed. As a result of this, the assessment techniques had to change to tie the assessment design and content to new outcomes. This integration of assessment with curriculum content and strategies is necessary to achieve the goals of assessment. Assessment and instruction should be inseparable if any educational program is to be successful (Marzano, Pickering and McTighe, 1993). The question is not whether to evaluate students, but how to measure performance in ways that would enrich learning, rather than restrict it. Effective assessment should be linked directly to instruction. It should be an essential part of teaching and every effort should be made to assure that what is measured flows from what is taught (Shrenko, 1994).

Reading is an essential skill as it is the vehicle for communication and sharing of ideas in authentic contexts in the global economy of the 21<sup>st</sup> century. In these contexts, students are expected to interpret, generate ideas of their own, make decisions, analyze, and evaluate others' ideas.

Despite the importance of EFL reading in this age, one of the immense problems facing public schools in Jordan is the low achievement of students in English language skills in general and in reading in particular. Jordanian secondary school students encounter many lexical, syntactic, and semantic difficulties when they are asked to read. They also cannot read what is between or beyond the lines.

From his 23-year experience as a teacher of EFL, the researcher claims that secondary school students cannot evaluate, analyze or interpret what they read. They cannot draw conclusions or generate innovative ideas, either.

The researcher believes that one of the causes of the previously-mentioned difficulties may be the reliance on traditional assessment in which reading texts are taught and assessed separately.

In support of the previously-mentioned probable cause, Wiggins (1993) states that traditional tests are "disrespectful" as they treat students as objects; as if their education and thought processes were similar and as if the reasons for their answers were irrelevant. In the same manner, Daniel (1999) mentions that traditional assessment assumes that all children have the same knowledge, and that limits the usefulness about individual student's learning styles or needs. He adds that educators need to offer up alternatives to traditional assessment.

One of the defects of traditional assessment according to Barth and Mitchell (1992) and Brooks and Brooks (1993) is that it does not show student understanding or evidence of student learning. Barth and Mitchell (1992) also mention that traditional tests corrupt teaching because students do not construct an answer. They further add that these tests give the impression that answers are always right or wrong and encourage memorization, rather than understanding. As a result, they "trivialize" schooling. According to Brooks and Brooks (1993), traditional assessment requires students to repeat back information they have received. They are not given the opportunity to demonstrate that they have constructed meaning about what they have learned. Consequently, many students struggle to understand concepts in isolation, to learn parts without seeing wholes and to make connections where they see only differences.

Hambleton, Jones, and Cadman (1993) also sum up some of the defects of traditional assessment. They mention that traditional tests foster a one correct answer mentally, narrow the curriculum, focus on the product at the expense of the process, require students to select answers rather than create them, focus on discrete skills, and encourage teaching to the test.

Falk (1994) reports the negative effect of traditional assessment on students' emotional state. She criticizes traditional testing and states that it is demoralizing for students. She explains that all the growth in self-esteem and self-confidence that takes place in a course of an entire year goes down the drain in the two or three hours of taking the test in its traditional format.

In searching for a solution to students' poor reading skills all over the world, many researchers (Arter and Spandel, 1992; Frazier and Paulson, 1992; Wolf, 1993;

Wiggins, 1993; Valencia, 1994; Eisner, 1999; Guskey, 2000; and Stiggins 2002) suggested using authentic assessment as an alternative to traditional assessment. In support of this suggestion, some researchers (e.g., Cross, Creer, and Pearce, 1998; Gagliano and Swiatek, 1999; Biondi, 2001; and Soyoun, 2003) found that authentic assessment improved students' poor reading skills. Therefore, this study attempts to improve secondary stage students' achievement in reading through the implementation of this type of assessment.

### Statement of the Problem

The aim of this study was to investigate the effect of an instructional program based on authentic assessment on the achievement of EFL secondary stage students in Jordan in reading.

### Elements of the Problem

The study aimed at answering the following question:

- Is there a difference in Jordanian Secondary school students' achievement in reading comprehension between the experimental group and the control group due to the method of assessment (Traditional or Authentic)?

### Research Hypothesis

The study examined the following hypothesis:

- There is no statistically significant difference ( $\alpha = 0.05$ ) in Jordanian secondary school students' achievement in reading comprehension between the experimental group and the control group due to the method of assessment (Traditional or Authentic).

### Significance of the Study

Recently, Jordan has witnessed a shift in promoting authentic assessment in schools. The researcher thinks that this research is very useful in exploring the effectiveness of authentic assessment in the EFL classroom and hopes to reach conclusions and suggest procedures on the appropriate practice to be used when applying authentic assessment in Jordanian schools. It may also help curriculum planners consider positively the usefulness of authentic assessment and, therefore, adopt it in upgrading EFL curriculum in Jordan.

### Definition of Terms

The terms below have the following definitions:

#### Authentic assessment

In the present study, this term refers to a type of assessment in which students demonstrate meaningful application of their reading skills as well as the strategies they employ while reading (Kunnan, 1998). This type of assessment includes many techniques such as interviews, dialogue journals, verbal reports,

conferences, learning logs, K-W-L-plus charts, anecdotal records, rubrics, self-assessment checklists and portfolios. The present study is confined only to the last three techniques (K-W-L-plus chart, self-assessment checklists and portfolios).

#### K-W-L-plus chart

In the present study, this term refers to a chart that students are requested to fill in with information before, and after reading. K stands for what the student knows about the title of the reading text, W stands for what the student wants to know, L stands for what the student learned from the reading text, and plus stands for what the student still wants to know.

#### Traditional assessment

In the present study, this term refers to discrete point assessment which is directed at discrete language components such as syntax, punctuation, spelling, lexical items and the like. This type of assessment includes multiple choice, true/false, short answer, filling in the blank tasks and the like.

### Review of Related Literature

#### Theoretical framework

Some of the benefits of authentic assessment are pointed out by Arter and Spandel (1992). They mention that authentic assessment displays critical thinking skills, problem solving, motivation and persistence in the learner.

One of the main benefits of authentic assessment according to Frazier and Paulson (1992) is that it is integrated with instruction. They explain that when authentic assessment is integrated in the learning environment, it gives students the opportunity to evaluate their work; thus, offering them a way to take charge of their learning. In this model, Wolf (1993) mentions that the authentic task that is used to measure the students' ability to apply the knowledge or skills is used as a tool for student learning.

Emphasizing that authentic assessment focuses on the authenticity of student learning, Wiggins (1993) states that authentic assessment refers to assessment tasks that resemble reading in the real world where students do not simply recall information or circle isolated vowel sounds in words; but apply what they know to new meaningful tasks. Brooks and Brooks (1993) also point out that authentic assessment is effective as the learning it measures has a value beyond the classroom.

Moreover, Valencia (1994) points out that authentic assessment develops strategic motivated readers; a goal which highlights the importance of a reader having knowledge of the reading process, knowledge about techniques and skills that are essential to constructing meaning or comprehending the text, and

knowledge about how to apply these techniques appropriately. In a similar vein, Eisner (1999) points out that as authentic assessment aims at moving away from testing practices that require students to select the single correct answer from four or five distracters, to a practice that requires students to create evidence through performance, it will enable assessors to make a valid judgment about what the student knows and can do in situations that really matter.

Authentic assessment can show students' different learning styles according to Geocaris and Ross (1999). They point out that authentic assessment allows the student to express knowledge in ways that best suit his/her individual learning styles and intelligence.

As reliability is one of the major concerns about any assessment method, Guskey (2000) emphasizes that authentic assessment is consistent and reliable. It involves teachers, students, and parents. In addition, it can indicate to a teacher what works and does not work in their instruction. It can allow teachers to analyze their personal strengths and weaknesses and initiate ways to improve the quality of their teaching.

Stiggins (2002) emphasizes that authentic assessment provides students with the opportunity to believe in themselves and to experience success in their learning environment. Success can lead to confidence which, in turn, can lead to increased effort and motivation.

To conclude, the previously-mentioned literature shows that authentic assessment assesses truly important aspects of learning such as higher order thinking skills and problem solving skills, and monitors growth and learning in social and emotional areas as well as academic areas.

### **Related Studies**

For Cross, Creer and Pearce (1998), Illinois grade-five students' scores indicated that traditional testing had failed to show accurately student knowledge in reading. Student reflection of the reading passages' content also proved to be inaccurate or shallow in displaying understanding of the literal and implied meanings. Using student portfolios, and authentic assessment techniques, Post- intervention data revealed that a majority of students scored higher when they were tested authentically in contrast with the traditional testing method.

Gagliano and Swiatek (1999) described a programme that was implemented to improve student achievement through the use of portfolios in language arts. Analysis of the probable cause data revealed that test scores did not provide a clear picture of student growth and development. Evidence of this problem was low test scores and poor quality work and differences between performance on tests and student ability. Post intervention data indicated that the use of portfolios

provided students with a sense of ownership in their own learning and with a more complete picture of student growth and development over time.

Biondi (2001) investigated the validity of authentic assessment techniques in a reading lesson. As students received practice on authentic assessment, their assessment became more aligned with the teacher's perceived assessment. The results of using self-assessment forms showed that 71% of the students acknowledged that their interest level in reading had risen. The researcher found that collaboratively-created checklists, simulations and portfolios in a reading lesson were valid, authentic and equitable measurement tools that would raise students' self-esteem and achievement through the assessment process.

Rhine and Smith (2001) carried out an 18-month program that focused on the problem of the inappropriate assessment of 539 primary grade students in reading comprehension. It was noticed that students at this stage showed low achievement while taking tests compared to their daily performance. The researchers used portfolios and checklists to assess the students. During the experiment, both traditional and performance tests were given and observed to show the difference in students' behaviour during testing situations. In the performance-based situation, the teachers realized that the student's fluency, sight word knowledge and comprehension improved. The researchers concluded that the students benefited from the authentic assessment strategies and had a greater understanding of their own learning.

Soyoung (2003) investigated the effect of authentic assessment on two college students' real performance and achievement. The learners were divided into two groups and their final assessment was assessed by their portfolios according to a grading rubric. After the application of the experiment, the final achievement scores of the control group and the treatment group were significantly different at 0.05.

In conclusion, these studies showed that using authentic assessment resulted in significantly higher reading achievement and better performance. These studies revealed that traditional assessment was ineffective as it provided information about a moment in time not the development of students' reading skills.

### **Design and Methodology**

#### **Subjects of the Study**

The researcher selected the subjects of the study from Sa'ad Bin Abi-Waqas Secondary School in Irbid First Directorate of Education. The sample consisted of two eleventh grade scientific stream. These classes were randomly assigned to an experimental group and a control group of 35 students each group.

## Instruments of the Study

To achieve the objectives of the study, the researcher used the following instruments:

(1) Two instructional programs as listed below:

- a. An Instructional program based on authentic assessment techniques.
- b. An Instructional program based on traditional assessment techniques.

(2) A reading achievement test.

### 1. The instructional programs

The researcher developed two instructional programs: one based on a) authentic assessment and the other based on b) traditional assessment. The instructional procedures were the same in both programs. However, the assessment techniques were different. Both programs were developed in light of the reading material included in *Jordan Opportunities 2* for the eleventh grade. In order to establish the validity of the two instructional programs, they were given to a jury of three supervisors and three instructors from the Department of Curricula and Instruction at Jordanian universities.

#### (a) The authentic assessment program

The authentic assessment program which was used with the experimental group students consisted of eight lessons. The techniques used in this program aimed at making students demonstrate meaningful application of their reading skills. The techniques that were employed were the K-W-L-S chart (see Appendix A) and the self assessment checklist (see Appendix B). The students were requested to fill in the K-W-L-S chart with information before, and after reading. The students were also requested to tick the strategies that they employed before, while, and after reading in the self-assessment checklist.

The reading assessment checklists and K-W-L-S charts were given to the students in every session and were collected by the teacher who later compiled them in a portfolio for each student. This portfolio was read by the teacher every week to diagnose the student's strengths and weaknesses.

#### (b) The traditional assessment program

The traditional assessment program which was used with the control group students consisted of eight lessons. The traditional assessment techniques which were used in this program were directed at discrete language components such as syntax, punctuation, spelling, lexical items and the like. They concentrated on the assessment of bits and pieces of information such as filling in blanks, sentence completion, matching, and the like. This information was mainly taken from the

Student's Book and the Activity Book of *Jordan Opportunities 2*.

### (2) The reading achievement test

The researcher developed a reading achievement test composed of ten questions about an expository passage. Five of these questions were literal, and the other five were inferential (see Appendix C). The questions of the test covered the main ideas in the text. The total score of the test was 10 points, 1 point for each correct answer.

#### (a) Validity of the reading achievement test

To ensure the validity of the reading achievement test, a jury of 5 EFL teachers, 3 EFL supervisors and 2 university professors was consulted, and their recommendations were taken into consideration before editing the final copy of this test.

#### (b) Reliability of the reading achievement test

To ensure the reliability of the reading achievement test, it was applied on a pilot study of 20 eleventh grade students which was excluded of the sample of the study. The test results were analyzed using the formula K-R 20 to correlate all items on this test with each other to make sure that all items were equivalent. The K-R 20 formula revealed that alpha was (.7835). This indicated that all the items of the test were equivalent.

### Study procedures

In this study, the researcher followed the following procedures:

- (1) Getting the approval of Irbid First Directorate of Education to conduct the study
- (2) Selecting the schools.
- (3) Selecting the subjects of the study from Irbid First Directorate of Education.
- (4) Pre-testing the experimental group and the control group to find out whether they were equivalent in reading. The means and standard deviations of the students' scores in the two groups on the reading pre-test were computed as shown in Table (1) below.

**Table (1):** The Means and the Standard Deviations of the Students' Scores on the Reading Pre-test

	GROUP	Mean	Std. Deviation	N
READ_PRE	experimental	4.94	2.511	35
	control	5.04	2.286	35
	Total	4.99	2.384	70

Table (1) shows that the mean of the experimental group students' scores in the reading pre-test was 4.94 and the standard deviation was 2.51, while the mean of the control group students' scores was 5.04 and the standard deviation was 2.28.

To determine if there was a statistically significant difference in the mean scores between the experimental group students and the control group students, a MANOVA analysis was carried out. The results of the

analysis of the pre-test scores are shown in Table (2) below.

**Table (2):** Results of the MANOVA for the Reading Pre-test

Source	Dependent Variable	Type III Sum of Squares	DF	Mean Square	F	Sig.
GROUP	READ PRE	.159	1	.159	.028	.869
Error	READ PRE	392.057	68	5.766		

Table (2) shows that the F value for reading was (.028). This value was not significant at 0.05. This indicates that the two groups were equivalent in reading before conducting the experiment.

(5) Developing the instructional program, the reading achievement test and ensuring the validity and reliability of the instruments of the study.

(6) Applying the programs from August to September 2009 by the researcher himself. During this month, the researcher trained the English teacher at the school to go on with the programs till November 2009.

(7) Post-testing the experimental group and the control group by the teacher to measure their achievement after conducting the experiment.

(8) Correcting the students' answers on the reading test by the teacher.

**Research Design**

The independent variables of this study were the assessment methods. These methods were the traditional assessment method and the authentic assessment method. The dependent variable was the students' achievement in the reading skills. The statistical design of this study was a quasi-experimental one. In this design, the researcher used an experimental group and a control group.

**Statistical analyses**

To analyze the data of this study, the researcher used means and standard deviations, the Multivariate Analysis of Variance (MANOVA), and the t-test.

**Findings of the study**

This section presents the findings of the study after analyzing the data. The findings are presented according to the hypothesis of the study.

The means and standard deviations of the students' scores in the two groups on the reading post-test were computed as shown in Table (3).

**Table (3):** The Means and the Standard Deviations of the Students' Scores on the Reading Post-test

GROUP	N	Mean	Std. Deviation
Experimental	35	7.09	1.716
Control	35	5.65	2.314
Total	70	6.37	2.148

Table (3) shows that there are differences between the mean scores of the students in the experimental and control groups. The mean of the experimental group students' scores is 7.09 and the standard deviation is 1.72, while the mean of the control group students' scores is 5.65 and the standard deviation is 2.31. The table shows that there is a difference of (1.44) between the means of the students' reading post-test scores in favour of the experimental group.

Therefore, t-test was used to determine the significance of the difference in the mean scores between the experimental and the control groups on the reading post-test. The results are shown in Table (4) below.

**Table (4):** The T- test Results for the Differences in the Mean Scores between the Experimental and the Control Groups on the Reading Post-test

GROUP	N	Mean	Std. Deviation	T	DF	Sig.
Experimental	35	7.09	1.716	2.954	68	.004
Control	35	5.65	2.314			

Table (4) shows that the difference in the mean scores of students' reading post-test was statistically significant (t =2.95, p =.004). This result shows that implementing authentic assessment was more effective in improving students' reading skills than traditional assessment. Therefore, the hypothesis of the study is rejected.

**Discussion and Recommendations**

The achievement of the experimental group students in reading could be attributed to the authentic assessment techniques which proved to be more effective than the traditional assessment techniques. This finding might be attributed to many reasons. First, authentic assessment techniques might have helped students discover their individual abilities and realize that there was not one right way to achieve success. In support of this reason, the researcher noticed that using the K-W-L-S chart helped students to build a bridge to understanding and enabled them to voice their thoughts, opinions and ideas. The researcher observed an increase in students' confidence and active participation in class

discussions because using the K-W-L-S chart encouraged brainstorming. This process promoted dialogue among students and raised their awareness of their new knowledge which allowed the researcher to tap into students' thoughts and helped them create the necessary connections for comprehension. Additionally, after correcting the experimental and the control group students' reading achievement test, the researcher computed the data on the two levels of the test, the inferential and the literal levels. On the inferential level, the analyzed data revealed that there was a significant difference between the experimental group students and the control group students' mean scores in favour of the experimental group students. However, there was no significant difference between both groups' mean scores on the literal level. The researcher realized that the self-assessment checklist might have attributed to this achievement on the inferential level as they provided the students with the higher-order skills needed for drawing conclusions and making judgments and inferences and helped them become independent and active learners. The K-W-L-S chart might also have had a contribution in enforcing the students' skills of prediction, speculation, and interpretation and their ability to express their ideas openly and freely without being criticized. Also in support of this, Cross et. al. (1998) found that students achieved higher scores when they were tested in an authentic way than when they were tested using traditional evaluation techniques. Additionally, Rhine and Smith (2001) found that student's knowledge and comprehension improved when authentic assessment was implemented.

Secondly, authentic assessment techniques might have increased students' self-confidence and raised their pride and feeling of accomplishment. This, in turn, could have improved their reading skills. In support of this reason, the researcher realized that the self-assessment checklist was not only effective in helping students self-evaluate, but it also helped them gain understanding of their strengths and the areas they need to improve. The self-assessment checklist also helped students to deal with more complicated assignments by employing the effective reading strategies that they learned in it. This might have made students more confident of their production and their ability to produce better work. It also might have made them proud as the amount and quality of work that was collected in their portfolios during the three months of the experiment was highly appreciated by the researcher. This result was supported by Biondi (2001) who found that self-assessment resulted in higher self-confidence, higher self-esteem and better achievement.

Thirdly, authentic assessment techniques could have helped teachers observe their students' reading skills and detect their strong points and their weak points through the use of portfolios. This might have enabled them to make plans to address areas that needed strengthening. In support of this, the researcher realized

that the experimental group students felt shy when their weaknesses were exposed at the beginning of the experiment. However, as the experiment went on, they started to accept these weaknesses and to discuss them more openly with the teacher. Thus, they started to show better performance in the reading classes and the researcher realized that their reading skills improved. By the end of the experiment, the portfolios could also show students' creativity and growth and they stressed the effort put forth by the students. The researcher felt that portfolios provided a multi-dimensional view of the students' interests, choices and learning styles and he also felt that he had gained insight into the true learning personality of each student. Also, in support of this reason, Soyung (2003) found that portfolios provided teachers with a more complete picture of the success of their teaching and areas where students needed improvement.

Fourthly, authentic assessment could have reduced students' tension which they might have felt during traditional assessment. This might have helped students learn better as nervousness and tension disappeared during authentic assessment. In support of this reason, the researcher realized that the self-assessment checklist took the mystery and ambiguity out of the assessment process. The students knew the criteria they were expected to meet when learning a new material ahead of time. Thus, tension and anxiety disappeared. The researcher realized that this anxiety-free atmosphere helped students improve their learning processes and consequently their reading skills.

### Recommendations

In light of the results of the study, the researcher recommended that the Ministry of Education take authentic assessment techniques into consideration during the process of designing English language curricula, that EFL teachers well-inform themselves about authentic assessment and incorporate it into their assessment practices, and that future researchers conduct further studies regarding the effect of using authentic assessment on other language skills such as listening and speaking.

### References

- Arter, J. and Spandel, V. (Spring 1992). Using Portfolios of Student Work in Instruction and Assessment. *Educational Measurement: Issues and Practices*, 33-44.
- Barth, P. and Mitchell, R. (1992). *Smart Start: Elementary Education for the 21st Century*. Golden Colorado: North American Press.
- Biondi, L. A. (2001). *Authentic Assessment Strategies in Fourth Grade*. Available at: <http://www.ERIC.ed.gov>. [ED 460 165].

- Brooks, J. G. and Brooks, M. G. (1993). *In Search of Understanding: the Case for Constructivist Classrooms*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Burke, K. (1999). *How to Assess Authentic Learning*. Arlington Heights, IL.: Skylight Publishing, Inc.
- Caustad, J. (1996). *Assessment and Evaluation in the Multiage Classroom*. Available at: <http://www.ERIC.ed.gov>. [ED 392149]
- Cross, B. J.; Creer, T.; and Pearce, M. (1998). *Improving Student Reading Comprehension skills Through the Use of authentic Assessment*. Chicago, Illinois: Saint Xavier University. IRI: Skylight Publishing Inc.
- Daniel, V. I. (1999). The Assessment Maze: Making instructional Decisions about Alternative Assessment for Students with Disabilities. *Preventing School Failure*, 43 (4) 171-178.
- Eisner, E. (1999). The Uses and Limits of Performance Assessment. *Phi Delta Kappan*, 80 (9), 658-660.
- El-Koumy, A. (2004). *Metacognition and Reading Comprehension: Current Trends in Theory and Research*. Available at: <http://www.ERIC.ed.gov>. [ED 490569]
- Falk, B. (1994). *Weaving Assessment into the Fabric of Teaching and Learning*. Available at: <http://www.ERIC.ed.gov> [ED 378217].
- Frazier, D. M. and Paulson, F. L. (1992). How portfolios motivate reluctant writers. *Educational Leadership*, 8, 62-65.
- Gagliano, K. and Swiatek, L. (1999). *Improving Student Assessment Through the Implementation of Portfolios in Language Arts*. Chicago, Illinois: Saint Xavier University and IRI Skylight Publishing inc.
- Geocaris, C.; and Ross, M. (September 1999). A Test Worth Taking. *Educational Leadership*, 29-33.
- Guskey, T. R. (2000). Twenty Questions? Twenty Tools for Better Teaching. *Principal Leadership*, 1 (3), 5-7.
- Hambleton, R. K.; Jones, R. W.; and Cadman, S. (1993). *Innovations in Testing and Evaluation of Student Competencies in Technical and Vocational Education*. ( Laboratory of Psychometric and Evaluative Research Report No. 230). Amherst, MA: School of Education, University of Massachusetts.
- Kohn, A. (September, 1999). Confusing Harder with Better: Why the "Tougher Standards" Movement is Undermining Our Schools. *Education Week*, XIX (2), 68.
- Kunnan, A. (1998). Language testing: Fundamentals. In A. Kunnan (ED.), *Validation in Language Assessment: Selected Papers from the 17<sup>th</sup> Language Testing Research Colloquium* (pp. 707-711). Mahwah, NJ: Lawrence Erlbaum.
- Marzano, R. J. ; Pickering, D. and McTighe, J. (1993). *Assessing Student Outcomes. Alexandria Association for Supervision and Curriculum Development*, v 1-5, 37-43.
- Meadows, S; and Karr, P. J. (2001). *The Role of Standardized Tests as a means of Assessment of Young Children*. Available at: <http://www.ERIC.ed.gov>. [ED 456134].
- Rhine, S. and Smith, E. (2001). *Appropriate Assessment of Primary Grade students*. Available at: <http://www.ERIC.ed.gov>. [ED 456 128]
- Rossi, R. (2000). Old-Style Teaching Gets Lower Scores. *Sun-Times*, (17).
- Shrenko, L. (1994). *Structuring a Learner-Centered Classroom*. Palantine: IRI/ Skylight Publishing Inc., 133-135.
- Soyoung, K. (2003). *The Effect of Authentic Assessment Strategy on Students' Achievement in a Constructivist Classroom*. Available at: <http://www.olam.ed.asu.edu>. Retrieved on February 19, 2007.
- Stiggins, R. (1999). Assessment, Student Confidence, and School Success. *Phi Delta Kappan*, 81 (3), 191-198.
- Valencia, S. W. (1994). *Authentic Reading Assessment: Practices and Possibilities*. DE: International Reading Association.
- Wiggins, G. (1993). *Assessing Student Performance: Exploring the Purpose and Limits of Testing*. San Francisco, California: Jossey-Base Publishers.
- Wolf, D. P. (1993). *What Works in Performance Assessment?* UCLA Los Angeles, CA: UCLA Graduate School of Education.