The Islamic University of Gaza

Deanery  of Graduate Studies

Faculty of Education

Curriculum & English Teaching Methods Department

# Evaluating English Writing Assessment in the 10<sup>th</sup> Grade at Gaza Schools with Regards to the Contemporary Trends

**Submitted by**

**Hosni Mohammad Sobeh**

**Supervised by**

**Prof. Dr. Ezzo Ismail Afanah**

**Dr. Nazmi Ihsan Al-Masri**

A Thesis Submitted in Partial Fulfillment of the Requirements for the

Master Degree of Education

**2007**

بسم الله الرحمن الرحيم

## الجامعة الإسلامية – غزة
## The Islamic University - Gaza

عمادة الدراسات العليا    هاتف داخلي: 1150

### نتيجة الحكم على أطروحة ماجستير

بناءً على موافقة عمادة الدراسات العليا بالجامعة الإسلامية بغزة على تشكيل لجنة الحكم على أطروحة الباحث/ *حسني محمد شحادة صبح* لنيل درجة الماجستير في كلية *التربية/* قسم *مناهج وأساليب تــدريس/* *اللغة الإنجليزية* وموضوعها:

## "Evaluating English Writing Assessment in the 10ᵗʰ Grade at Gaza Schools with Regards to the Contemporary Trends"

وبعد المناقشة العلنية التي تمت اليوم السبت 13 رمضان 1429هـــ، الموافــق 2008/09/13م الـساعة الثانية عشرة ظهراً، اجتمعت لجنة الحكم على الأطروحة والمكونة من:

| | | |
|---|---|---|
| أ.د. عـزو عفـانــة | مشرفاً ورئيسـاً | ........................... |
| د. نظـمي المصري | مشــــرفـاً | ........................... |
| د. عبد المعطي الأغا | مناقشاً داخليـاً | ........................... |
| د. بـاسـل سـكيـك | مناقشاً خارجياً | ........................... |

وبعد المداولة أوصت اللجنة بمنح الباحث درجة الماجستير في كلية *التربية/*قسم *مناهج وأساليب تدريس/* *اللغة الإنجليزية.*

*واللجنة إذ تمنحه هذه الدرجة فإنها توصيه بتقوى الله ولزوم طاعته وأن يسخر علمه في خدمة دينه ووطنه.*

### والله ولي التوفيق ،،،

عميد الدراسات العليا

د. زيــاد إبراهيم مقداد

# Abstract

This descriptive analytic study aims at evaluating English writing assessment in the 10th grade at Govermental Gaza schools with regards to the contemporary trends in the scholastic year 2006-2007. It sought to determine the existence of the contemporary trends in assessing the 10th grade students' writing in Gaza schools.

Based on the previous studies, literature review and the theoretical framework, the researcher designed a questionnaire to investigate the extent to which the contemporary methods of assessment are used by the 10$^{th}$ grade teachers.

The questionnaire involved 10 assessment methods as follows: Testing, observation, portfolios, self-assessment, peer-assessment, journals, interviews and conferences, rubrics, anecdotal records and checklists. Validity and reliability were achieved through refereeing the questionnaire by a panel of experts and the statistical procedures.

The traditional methods of assessment, testing and observation got the highest score in a percentage of 63.8% for testing followed by observation in  a percentage of 62.8%. The rest of the eight method which were classified as the contemporary methods got low percentages .  Their  percentage ranged between 28.7% " to 22.4%. The results indicated that most teachers still rely on the traditional methods in assessing their students' writing.

Recommendation were drawn to adopt the contemporary methods of assessment that suit our context in Gaza. That should comprise all levels of the education system including the Ministry of Education , supervisors, principals and the teachers.

:

.2007-2006

.

:

.

.

.  62.8                        63.8

.        22.4              28.7

.

# Dedications

*This thesis is dedicated:*

To the soul of martyrs, who sacrificed their lives for Al-Aqsa Mosque and Palestine.

To the soul of my father who waited my success but he passed away before achieving my study.

To my mother for her endless patience and unwavering support that she has shown to me during this long, arduous process. She has been my largest source of inspiration, empowers my career of education and waits my success.

To my lovely wife, my sons Khaled, Abdul-Aziz and Ahmed, and my daughters Bayan and Iman, for their endless patience and support that they have shown to me during this process.

To all with warm regards.

# Acknowledgments

I am extremely grateful to my supervisors and committee chair Prof. Dr. Ezzo Ismail Afanah and Dr. Nazmi Ihsan Al-Masri for their support, wisdom, and encouragement along the way to keep going.

To Dr. Nazmi Al-Masri: Your guidance, encouragement, and advice have made this work possible. Thanks for the unique consultations in my research and effort during my MA study. I deeply appreciate your advice, guidance, and time

I am so grateful to all my teachers who contribute their time and effort to facilitate my learning and research.

I would like to thank Drs Akram Habib, Walid Amer, Kamal Murtaja and Awad Keshta for their unique consultations on my research and efforts during my MA dissertation.

Special applause to my colleagues Mr. Mohammad Al-Mazloum, Yousif Hamdona and Mahmoud Mattar who devoted their time and experience to cooperate with me during my dissertation.

Many thanks are extended to all those who completed the study with me.

# Table of Contents

| | |
|---|---|
| **Conclusion and recommendations** | |
| Introduction | 117 |
| Discussion | 117 |
| **Suggestions** | 124 |
| Recommendations for further studies | 128 |
| References | 129 |
| Appendices | |
| **List of Tables** | |
| Table (1) Differences between traditional and alternative methods of assessment (Anderson, 1998) | 18 |
| Table (2) A sample of self-assessment in writing | 23 |
| Table(3) Checklist to assess a student's handwriting by: Gomez et al (2005) | 33 |
| Table (4) An example of an anecdotal record by Batstone (2004) | 34 |
| Table (5) An example of a rubric used to assess an article | 37 |
| Table (6) The Population of the Study | 73 |
| Table (7) The Sample of the Study | 74 |
| Table ( 8) The correlation coefficients between the Tool & its domains | 77 |
| Table (9) Discriminate validity | 78 |
| Table (10) Cronbach Alpha Reliability Coefficients | 79 |
| Table ( 11 ) Split Half Method: | 80 |
| Table ( 12 ) Descriptive Statistics for Testing Domain | 90 |
| Table(13) Reasons for not using testing in assessment | 90 |
| Table (14 ) Descriptive Statistics for Observation Domain | 91 |
| Table (15) Reasons for not using observation in assessment | 92 |
| Table (16 ) Descriptive Statistics for Portfolio Domain | 93 |
| Table (17) Reasons for not using observation in assessment | 93 |
| Table (18 ) Descriptive Statistics for Self assessment Domain | 95 |
| Table (19) Reasons for not using self- assessment method in assessment | 96 |
| Table (20 ) Descriptive Statistics for Peer assessment Domain | 97 |
| Table (21 ) Reasons for not using Peer assessment in assessment | 98 |
| Table ( 22 ) Descriptive Statistics for Journals Domain | 99 |
| Table ( 23 ) Reasons for not using Journals in assessment | 100 |

# Chapter I

1. Introduction

2. Statement of the problem

3. Questions of the study

4. Purposes of the study

5. Significance of the study

6. Limitations of the study

7. Procedures of the study

8. Definitions of terms

# Chapter I

## 1. Introduction

There has been a recent growing interest among educators, policymakers in alternative assessment due to concerns that traditional types of assessment do not assess higher order skills and other skills essential for functioning in school or work settings. Traditional tests cannot be used to closely monitor student progress in the school curriculum throughout the year since they are only administered once or twice annually. Students are put under too much pressure, their want to learn is damaged. Feedback may be limited to  marks or grades and students may not have the opportunity to make sense of the feedback they receive. Furthermore, students may not be aware of the criteria used to assess their work and learning may become driven by assessment, and students may only do those things that are assessed.

Effective methods of assessment should be part of instruction and learning, and provide reflection from students as well as teachers. The movement away from traditional tests to alternative assessments is called authentic assessment or performance assessment. These terms and assessment strategies have led the quest for more meaningful assessment which better capture the significant outcomes we want students to achieve and better match the kinds of tasks which they will need to accomplish in order to ensure their future success.

The integration of assessment with curriculum content and strategies is necessary to achieve the goal of assessment. Assessment and instruction must be inseparable if the teaching program is to be successful (Marzano, Pickering, & McTighe, 1993). Further, depending on tests as the sole means of assessment  lead to narrowing the school curriculum by directing teachers to focus only on those subjects

and skills that are included in the examinations. Consequently, such tests may dominate and distort the whole curriculum.

Learner-centered classrooms and alternative assessments go hand in hand. In a learner-centered class, the assessment system assesses different students differently, and includes student input in design and revision. It monitors progress continually in order to provide feedback on individual growth and progress. Students can choose the appropriate types of products for demonstrating achievement of educational standards. It can further promote students' reflection on their own growth as learners through opportunities for self-assessment, and allow diversity of competencies to be demonstrated in a variety of ways.

At the  psychological level, it is believed that the role of the students in the contexts where tests are introduced is that of passive recipients of knowledge and their needs and intentions are generally ignored. Traditional tests are also said to have detrimental consequences on students' intrinsic motivation, self-confidence, effort, interest and involvement in the language learning experience and induce negative feelings in students such as anxiety, boredom, worry and fear, which are not conducive to learning  (Tsagari, 2004).

Furthermore, alternative assessment is also said to be in line with views expressed in cognitive psychology, which suggest that learning is not linear, but proceeds in many directions at once and at an uneven pace. (Ibid) The focus of traditional tests is usually on small pieces of information, but using alternative methods of assessment make learners know how to organise, structure, problem-solve, experiment, make decisions, cooperates with others and produce a product. These situations simulate real-world activities (Borich, 2004: 475).

Dietel et al (991 in Tsagari, 2004) stress that alternative assessment techniques allow learners plenty of time to 'generate' rather than 'choose' a response: after recently acquired knowledge is brought to the forefront of their minds, the higher-order thinking skills of synthesis and analysis are required for the learners when participating in alternative assessment activities, which they can later reconsider by critically working together with the teacher or other learners in sharing perceptions.

Genesee and Upshur (1996) stress that alternative assessment methods can also gather information about those factors that affect student achievement which should be seen as an integral part of students' assessment

• learning strategies (e.g. whether the student takes risks, improvises, focuses on meaning/form, self-corrects, uses first language strategies)

• affective and personality styles (e.g. whether the student is enthusiastic, self-reliant, resourceful, passive)

• students' work habits (e.g. whether the student is punctual, follows instructions well, meets goals, prepares for class homework, seeks assistance when needed)

• students' social behaviour (e.g. whether the student works cooperatively, socialises with peers, participates in class discussion)

• reactions to the course (e.g. student participates actively in class activities, requires extra guidance, shows initiatives.

The Ministry of Education and Higher Education (MEHE)  has implemented the first Palestinian curriculum plan by developing textbooks for both the basic and secondary stages. The first textbook was produced for the first basic primary grade one in 2000. Grade Ten textbook was produced in 2004 and taught in the scholastic year 2004 – 2005. In spite of the developing of the textbooks, teachers in Gaza schools still depend on traditional tests as their major means of assessment. They

stress using tests and term papers as their main resources for assessing student work. Consequently, tests are used just for grading students and no role is considered to assessment as a contributor to the development of the learning process.

This study offers alternative methods for writing assessment in order to develop writing among the Palestinian tenth grade students in Gaza. It would also help schools to develop an effective assessment plan that may yield meaningful data at the student, course, program, and school levels.

## 2. Statement of the problem

Current trends in assessment express dissatisfaction with testing; the traditional method of assessment, which is faulted for not capturing vital information about students' competence in their second language. Accordingly, a widespread call emerges for increasing the use of meaningful alternatives forms of assessments that involve language students in selecting and reflecting on their learning and in which language teachers have a wider range of evidence on which to judge whether students are becoming competent and purposeful language users. Additionally, these alternative forms of assessment are believed to instill in students lifelong skills related to critical thinking that build a basis for future learning, and enable them to evaluate what they learn both in and outside of the language class.

However, current instruction in the English program at governmental or UNRWA schools in Gaza lacks these alternative forms of assessment. As a result, the teaching and learning process continues to rely heavily on traditional forms of assessment.

### 3. Questions of the study

The study attempts to answer 5 questions to achieve its purpose

1. What are the current trends for developing the assessment of writing in English language?

2. What are the existing methods used by teachers for assessing English writing in 10th grade at Gaza schools from the teachers' perspective?

3. To what extent do the existing methods for assessing writing in English in 10th grade at Gaza schools match those of current trends?

4. Is there a significant statistical difference at $(\alpha \leq 0.05)$ related to the assessment methods used by male and female teachers?

5. Is there significant statistical differences at $(\alpha \leq 0.05)$ related to the assessment methods used by teachers according to their experience?


### 4. Purposes of the study

1. To identify the current methods used for assessing English writing in 10th grade at Gaza schools and their pedagogical benefits.

2. To identify the existing methods used for assessing English writing in 10th grade at Gaza schools.

3. To explain the extent of which these current methods used for assessing writing in English in 10th grade at Gaza schools match the modern methods for assessing English writing .

### 5. Significance of the study

The current study is significant because it:

1.  provides teachers with the current methods used for assessing writing in English   language tests.

2.  finds out the methods used by teachers to assess writing in the 10th   grade English  language writing at Gaza schools .

3.  contributes to the existing body of literature about language assessment   .

### 6. Limitations of the Study

The study is applied in accordance with these limitations:

1.  This study  is restricted to the English learners of  Gaza schools  and cannot be generalized beyond.

2.  It intends to cover  English learners in the tenth grade.

3.  This research intends to cover writing skills and does  not include other language skills

### 7. Procedures of the study

The study adopts the following procedures to accomplish its objectives:

1.  Examining previous studies to help the researcher benefit from their procedures, tools, results, and recommendations.

2.  Preparing theoretical framework through reviewing the literature concerned.

3.  Designing a questionnaire for the purpose of the study.

4.  Consulting experts in English language methodology to ensure validity and reliability of the questionnaire.

5.  Analyzing the collected data and giving interpretation.

6.  Giving recommendations in the light of the study findings.

## 8. Definition of terms

### 8.1. Assessment

Assessment is defined as data-gathering strategies, analysis, and reporting processes that provide information which can be used to determine whether or not intended outcomes are being achieved (Foundation Coalition, 2007).

### 8.2. Evaluation

It is a wide process of interpreting data to make judgments about a particular programme. It includes collecting, analyzing and interpreting information about teaching and learning in order to make informed decisions that enhance student achievement and the success of educational programmes ( Allen, 1998).

### 8.3. Alternative assessment

Alternative methods of assessment include every method or procedure used to assess learners which are not normal tests or standardized exams. There are several alternative methods like: observation, portfolios, rubrics, self-assessment, peer-assessment, checklists, journals, interviews, conferences and anecdotal records (Lafi, 2002).

### 8.4. Testing:

Brown (2004:3) defines tests as methods of measuring a person's ability, knowledge, or performance in a given domain. Most common forms of tests include fill–in–the blanks, sentence completion, open answers, and multiple choice.

## 8.5. Higher Order Thinking Skills (HOTS):

Watson (2003) defines HOTS as the moving away from general knowledge type skills to thinking skills like: synthesizing, analyzing, reasoning, comprehending, application, evaluation. The emphasis is shifted from drill and repetition activities to problem solving and higher level/order thinking skills.

## 8.6. Traditional assessment

According to many researchers, tests are the only traditional type of assessment. Many teachers and educators resorted, in the past and still in some areas, to tests as the basic means for assessing students because they are easy to administer and score and the data obtained are easier to interpret when comparing performance across groups of students or across time. (Shaaban, 2001 and Lafi, 2002)

## 8.7. Writing:

System of human communication using signs or symbols associated by convention with units of language — meanings or sounds — and recorded on materials such as paper, stone, or clay. (Britannica Online Encyclopedia)

## 8.8. A product Writing

This is a traditional approach, in which students are encouraged to mimic a model text, which is usually presented and analysed at an early stage (Greece, 2002)

### 8.9. A process Writing

The process writing refers to a teaching approach that focuses on the process a writer engages in when constructing meaning. This includes different stages of the writing process such as: pre-writing, writing and re-writing.(Greece, 2002)

### 8.10.  EFL

EFL refers to English as a foreign language and is used to refer to English instruction that occurs in a non-English-speaking context such as in Palestine, where the language of mass communication is Arabic (Echevarria et.al, 2004:221 ).

### 8. 11. TESOL

It refers to teaching English to speakers of other languages and the training of teachers in the methodologies of English as a second or foreign language instruction (Echevarria et.al, 2004:222 ).

### 8. 12.  10th   Grade Students in Gaza

The tenth grad students are the student who live in  Gaza and enrolled in the governmental schools under the control of the Palestinian National Authority. The majority of student suffer poverty as a result of the siege.  The number of households in Gaza below the poverty line has reached nearly 52 percent, according to a report in 2007   by   the   United   Nations   Relief   and   Works   Agency   (UNRWA)   .

# Chapter II

# Literature Review

1. Section I  Theoretical Framework

2. Introduction

3. Testing

4. Evaluation

5. assessment

6. Alternative assessment

7. The Rationale for the Shift to the Alternative Modes of Assessment

8. Characteristics of alternative assessment

9. Types of alternative assessment

10. Conclusion

11. Section II  Previous studies

12. Conclusion

# Chapter II

## Section I        Theoretical framework

## 1. Introduction

Recent years have witnessed a move away from the traditional assessment methods to alternative methods of assessment. These alternative methods show what the students can actually do rather than what they are able to recall and provide multiple ways of determining the progress of students**.**

Genesee and Upshur (1996: 23) state that decisions about instruction require more than data on students achievement. They require information about students' needs, goals and attitudes towards school and learning and this cannot be obtained from any single assessment procedure.

> Traditional, so-called standardized testing methods representative of the testing culture, such as multiple-choice tests, true/false items or short-answer tests, were increasingly criticized for not being suitable for the changed educational goals aiming at competency development. (Moskal, 2003)

Traditional types of assessment were considered inadequate for measuring higher-order thinking skills  and were seen as stimulating students to adopt surface study strategies such as memorization and reproduction at the expense of deep study activities. In contrast alternative methods of assessment stimulate students to integrate knowledge, skills and attitudes and use them to solve realistic professional tasks (Birenbaum, 1996; Dochy, 2001; Reeves & Okey, 1996).

In this chapter, an explanation  is made for different terms: testing , assessment and evaluation. In addition to the traditional tool, testing, the researcher identifies nine tools for alternative assessment with their definition and some of their characteristics.

The tools included are observation, self- assessment, peer-assessment, journals, logs, anecdotal records, conferences and interviews, portfolios, and rubrics.

**2. Testing**

Tests are the most familiar example of assessment and the most widely used tool in the classroom. Most common forms of tests include fill–in–the blanks, sentence completion, open answers, and multiple choice. According to Allan (1995), tests are instruments of evaluation by which we try to measure learner performance. They have a physical existence and operate within specific time frames, seeking to make accurate predictions on the basis of relatively small samples of performance in the case of such an enormously complex thing as language.

Some advantages of tests lie in their capacity to provide a quick and easy way to tell what information has been obtained, what broad skill level a person may be functioning on, what vocabulary has been learned and what sub–skills have been acquired.

Some disadvantages of using tests such as: short-answer or selected-answer assessments was narrowing of the curriculum by directing teachers to focus only on those subjects and skills that are included in the examinations. Teachers study the tests to see what is being assessed since they, as well as the students, are being held accountable for the test results. So many teachers emphasize what the tests cover and model instruction after them. This narrowing of the curriculum led some textbook authors and publishers to structure textbooks and instructional materials that reflect the content and skills emphasized on the tests. Such textbooks and other materials provide learning activities that mimic what the tests have asked students to do. The

published materials and teachers have led students to practise isolated objectives and fractured skills which are usually emphasized on the tests (Farr & Tone,1994).

In many schools, raising test scores has become the single most important indicator of school improvement. As a result, teachers and administrators feel enormous pressure to ensure that test scores go up. Schools narrow and change the curriculum to match the test. Teachers teach only what is covered on the test. Methods of teaching conform to the multiple-choice format of the tests. Teaching more and more resembles testing.

Teaching to the test also narrows the curriculum by forcing teachers and students to concentrate on memorization of isolated facts, instead of developing fundamental and higher order abilities. For example, multiple-choice writing tests do not measure the ability to organize or communicate ideas. (The National Center for Fair & Open Testing, 2006)

Many students pass through the experience that they study hard, participate in class, do all of their homework, and they think they have a grip on the material. But then the day of the test comes. Suddenly, they blank out, freeze up, lose concentration, or feel so nervous that they cannot respond to those questions they knew the answers to just last night. The students may also be facing extenuating circumstances (e.g., personal problems or illness) at the time they are being tested, thus also hampering their performance on the test ( Macias, in Richards and Renandya, 2002: 338).

Negative effects on teachers' psychology are reported due to the dictate of tests to reduce the professional knowledge and status of teachers and exercise a great deal of pressure on them to improve test scores which eventually makes teachers

experience negative feelings of shame, embarrassment, guilt, anxiety and anger. (Tsagari, 2004)

## 3. Assessment

Assessment is defined as data-gathering strategies, analyses, and reporting processes that provide information which can be used to determine whether or not intended outcomes are being achieved (Foundation Coalition, 2007*).*

Lake (1998) stated that assessment is a process of gathering information to meet a broad range of evaluation needs. The primary purpose of assessment is for the student to receive multiple attempts to practice and to demonstrate understanding of content and to develop skills by receiving specific and timely feedback by the teacher in order to improve achievement. The primary purpose of assessment is for the teacher to analyze student progress for the purpose of modifying and refining the teaching/learning cycle to better meet student needs.

Assessment differs from testing in that the former is a continuous, on-going process while the latter is limited in a specific time. The purpose of assessment is to provide feedback to the learner and the teacher about strengths and weaknesses during the learning process and to encourage constant improvement.

## 4. Evaluation

Evaluation is broader in scope than testing or assessment. It is a wide process of interpreting data to make judgments about a particular programme. It includes collecting, analyzing and interpreting information about teaching and learning in order to make informed decisions that enhance student achievement and the success of educational programmes (Lafi 2002).

The focus in evaluation can be on different aspects of teaching and learning like: textbooks and instructional materials, students' achievement and whole programmes of instruction. As stated by the American Evaluation Association (2007), evaluation involves assessing the strengths and weaknesses of programs, policies, personnel, products, and organizations to improve their effectiveness.

According to Genesee and Upshur (1996: 5), the overall purpose of second language evaluation is to make sound choices that will improve second language teaching and enhance second language learning. Three essential component of evaluation are information, interpretation and decision making.

The relation between evaluation, testing and assessment can be summarized as follows: Evaluation is the umbrella that encompasses testing and assessment. It goes beyond student achievement and language assessment to consider all aspects of teaching and learning and to look at how educational decisions can be informed by the results of alternative modes of assessment.

## 5. Alternative assessment

The realization of the importance of assessment to student learning resulted in reforms in assessment that shift the emphasis from the traditional measurement to alternative methods of assessment. Alternative methods of assessment include every method or procedure used to assess learners excluding normal tests or standardized exams. Alternative Assessment applies to any and all assessments that require students to demonstrate knowledge and skills in ways other than through the conventional-methods used within a classroom. Pierce and O'Malley (1992) consider alternative assessment as "any method of finding out what a student knows or can do that is

intended to show growth and inform instruction and is not a standardized traditional test."More details about alternative assessment are discussed below.

**5.1. The Rationale for the Shift to the Alternative Modes of Assessment**

The dissatisfaction with traditional testing techniques led to the search for new alternative modes of assessment . The following are some shortages of traditional assessment which question their effectiveness for being suitable methods of assessing writing in our context.

> 1-Traditional ways of testing can sample only a small part of what we want to produce.
> 2-Traditional tests play a judgmental role, not a developmental one: they are used to make judgments about success or failure, to select or exclude . They have little to do with improving the overall quality of language learning and teaching.
> 3-Traditional test are summative and usually designed and administered by people outside the learning process.
> 4-Tests take up a lot of teaching time and are very often not followed up as a basis for future teaching.
> 5-Tests tend to treat learners as powerless victims rather than active participants in the learning process.
> 6-Tests are teacher-centered.
> 7-Testing has limited time frames; it is a "one shot" event that gives the learner only one chance to show competence.
> 8-The test results may not be a true reflection of what a learner can do; they can tell us about certain aspects of student achievement. They cannot tell us much about the other factors that often figure in foreign language evaluation.
> 9-Testing is anxiety-generating, and this may affect the test-takers' learning as well as their self-image. They may also be very demotivating.
> 10-Tests are administered to large groups of students. They are not individualized and cannot be tailored to the needs of individual learners.
> 11-Tests are not always fair as they do not account for individual differences, multiple intelligences, different learning styles (Lafi, 2002)

It can be concluded that testing is a means to assess the learner, an instrument of evaluation used to measure the learner's performance. Tests are meant to give us a sample of the big picture. They measure only  a few aspects of skill development.

They cannot be considered as the single indicator of school improvement but one variable of the many methods of assessment that are available.


## 5.2. Characteristics of alternative assessment

The search for more valid and reliable forms of assessment leaded to alternative methods of assessment that closely match today's educational goals and capture evidence of best performance across time. "Rationales such as constructivism, obtaining student generated responses, thinking of learners as individuals, and linking assessment to instruction have been used to support the use of alternative assessments (Veronesi, 1997).

Macias, in (Richards and Renandya, 2002:339) shows several characteristics of alternative assessment:

1.  They do not intrude in regular classroom activities .

2.  Reflect the curriculum that is actually being implanted in class.

3.  Provide information about both the strengths and weaknesses of students.

4.  They are multi-culturally sensitive and free from, linguistic ,and cultural biases found in traditional testing.

5.  Provide multiple indices that can be used to gauge student progress.

Lafi (2002) offers more characteristics to alternative assessment:

1.  They are in congruence with the learner-centred principles since they view the learner as an active agent in the assessment process rather than a powerless victim of the testing techniques.

2.  They treat assessment as an integral part of teaching and learning processes, which are closely related to instruction.

3. They take into account variations in students' needs, interests and learning styles, i.e. they can be individualized: students do not learn in the same way, they cannot, as a result, be assessed in the same fashion.

4. They are on-going and carried out overtime. They provide the learners with opportunities to revise, improve, add, etc…

5. They can help decrease the level of anxiety and increases students' comfort zone and feeling of success.

The following table outlines differences between traditional and alternative methods of assessment .

**Table (1)**

**Differences between traditional and alternative methods of assessment**

**(Anderson, 1998)**

| TRADITIONAL ASSESSMENT | FEATURE | ALTERNATIVE ASSESSMENT |
|---|---|---|
| Universal assessment | Knowledge | Multiple Meanings |
| Passive process | Learning | Active Process |
| Process and product are separate | Process | Process and product are emphasized together |
| Information in discrete bits | Focus | Inquiry: the nature of questioning |
| Documentation of learning | Purpose | To facilitate learning |
| Cognitive, affective, and conative abilities are separate | Abilities | Cognitive, affective, and conative abilities are connected |
| Assessment is objective and value neutral | Assessment | Assessment is subjective and value-laden |
| Hierarchical model | Power and Control | Shared between teacher and learner |
| Learning is an individual process | Nature of Learning | Learning is a collaborative process |

**5.3. Types of alternative assessment**

Alternative assessment includes a variety of instruments that can be adopted or adapted to a variety of situations. No single instrument will fit the needs of a given group of students, thus the idea is to adapt and adopt the most suitable instrument and to implement this instrument in a way that reflects the goals of the class**.**

There is no agreement among educators and researchers on specific methods of alternative assessment. Teachers need to determine which method or combination of methods is the most appropriate for collecting information that can help them to assess their students. All of these methods can reveal important data about student learning and effective instruction. Some of these methods include observation of student behavior through lessons, comments by students during conferences or from students' journals. Other methods like portfolios and conferences give students opportunities to use language with teachers in ways that rarely occur during class time. They give students a sense of involvement and increase their enthusiasm for learning.

The following methods of assessment are the most prominent in literature. All of them are considered alternative methods for assessing students writing

**5.3.1. Observation**

Informal observation is an integral part of everyday teaching; indeed, teachers continuously observe their students' language use during formal instruction or while the students are working individually at their desks. "Over periods of time they have the opportunity to observe learners participate in a wide range of activities and tasks, working on their own and in groups, developing their ability to communicate with others" (Flucher and Davidson ,2007 : 25).

Teachers can then detect changes in student achievement and make decisions about what should follow. Teachers should record the results of their observations to keep track of individual students on the whole group's achievement, progress, difficulties, strengths, etc, and think of remedial work if need be.

### 5.3.1.1. Types of Observation

Maxwell (1999) classified teacher observation into two types: incidental and planned.

1. Incidental observation occurs during the ongoing (deliberate) activities of teaching and learning and the interactions between teacher and students. In other words, an unplanned opportunity emerges, in the context of classroom activities, where the teacher observes some aspect of individual student learning. Whether incidental observation can be used as a basis for formal assessment and reporting may depend on the records that are kept.

2. Planned observation involves deliberate planning of an opportunity for the teacher to observe specific learning outcomes. This planned opportunity may occur in the context of regular classroom activities or may occur through the setting of an assessment task (such as a practical or performance activity).

Maxwell (1999) stated that using teacher observation for assessment has a strong justification due to its capacity to enhance assessment validity. By extending the range of possible assessments, teacher observation allows assessment to be more:

1. comprehensive — ensuring recognition of all desired learning outcomes, especially those not otherwise assessable than in classroom contexts;

2. connected — situated within familiar learning contexts and closely related to curriculum frameworks, learning experiences and pedagogical planning;

3. contextualized — sensitive to the effects of context on performance and deriving assessment evidence from a variety of situations and occasions;

4. authentic — interesting, challenging, worthwhile and meaningful to students;

5. holistic — emphasizing relatedness and connections in learning and involving performance on complex wholes rather than separate components.


**5.3.1.2. Characteristics of Observation**

Using observation as a method of assessment has many advantages for teachers by informing them about learners and the effectiveness of their teaching. Observation allows teacher to identify some learning strategies used by learners. For instance, regarding meta-cognitive strategies, teachers can notice whether learners have prepared for class or not, whether they keep a journal or vocabulary book, or whether they self-correct. In cognitive strategies, repetition, resourcing and inferencing are all identifiable in class and where the teacher knows the students' L1, translation is also often evident. Finally, the degree to which students use socio-affective strategies such as cooperation and asking for clarification or confirmation can be determined by observing students' willingness to participate in pair and group work in class or other activities which require interpersonal contact. (O'Malley, 1992)

Observation was considered more accurate than student self-assessment because previous work with students showed that what they thought about themselves and actually what they did in class did not always closely correspond (Vose, 1997).

**5.3.2. Self-Assessment**

Self-assessment is to involve students in observing, analyzing, and judging their performance on the basis of criteria and determining how they can improve it. Assessment decisions can be made by students on their own essays, reports, projects, presentations and even exams . To be able to perform self-assessment, learners need preparation and guidance and a great deal of practice in order to do it successfully and feel comfortable about it. They must also be able to understand and select suitable criteria on which to base the assessment and choose a realistic standard against which to measure achievement (Alverno College Faculty, 2006).

In the communicative approach, the teacher is described as a facilitator who should create conditions to foster learner independence and develop the learners' "how-to-learn" skills. The teacher is expected to provide the proper conditions and devise the appropriate tools to train the learners to assess themselves, reflect on their learning experience and identify their strengths and weaknesses. (Coombe, 2002).

**5.3.2.1. Characteristics of self-assessment**

Some advantages of self-assessment are that it allows the learners to monitor his own performance in a stress-free setting. It also gives them the feeling of responsibility, self-reliance and this in turn may enhance motivation.

Self –assessment is extremely important in that it promotes invaluable learning skills such as monitoring one's own progress, reflecting on one's abilities and learning styles, and setting personal goals .( Georgiou & Pavalou,, 2003 : 10)

Isaacs (1999) mentioned that using of self assessment helps students to examine their own work critically and provides some feedback to students on their work without

imposing too heavy an additional burden on the teaching staff. Besides, it is a way of ascribing a mark or grade to a student's work for summative purposes.

One of the most obvious benefits of self assessment according to East (2006) is that it relates very closely to the aims of personal development planning. This involves students engaging in critical self reflection, focusing on obtaining a clearer idea of the features of effective learning and thereby increasing their understanding of the subject matter being studied. Allan et. al (1998) show a sample of self – assessment can be used in writing.

**Table (2)**

**A sample of self-assessment in writing by Allan et. al (1998)**

| |
|---|
| Name:                                                   Date: |
| General remarks regarding my creative writing (e.g., genres and topics presently working on): |
| Things I do well: |
| Areas where I have shown recent improvement: |
| Areas needing further work: |
| My goal for my creative writing this term: |
| Steps I will take to attain my goal: |
| Literary works I might study to help me solve problems I am encountering: |
| Others concerns or comments: |

### 5.3.3. Peer-Assessment

Peer assessment is the process whereby learners respond to and evaluate each other's performance or achievement. It can promote higher order thinking and support cooperative learning. It can take many forms: the spoken word, the written word, checklists, non-verbal symbols, numbers along a scale, etc.

Students need training in assessing work, especially if they are encountering peer assessment for the first time. They need to have the chance to learn and practice the relevant skills of applying criteria, giving feedback, and applying standards and deriving grades. Students can benefit from using rubrics or checklists to guide their assessments. At first these can be provided by the teacher; once the students have more experience, they can develop them themselves. (NCLRC, 2004)

### 5.3.3.1. Characteristics of peer assessment

Some reported advantages of using peer assessment dwell in the following aspects: Peer assessment supports students to become more autonomous, responsible and involved. It helps to consolidate, reinforce and deepen understanding by engaging students in cognitively demanding tasks: reviewing, summarizing, clarifying, giving feedback, diagnosing misconceptions, identifying missing knowledge, and considering deviations from the ideal (Hamer et.al, 2005).

In addition, students are given a wider range of feedback. Even if the quality of feedback is lower than from professional staff, its immediacy, frequency and volume may compensate (NCLRC, 2004).

"During the peer assessment process, students are exposed to a variety of styles, techniques, ideas and abilities, in a scale of quality from mistakes to exemplars." (Hamer, et.al , 2005) .

Other important advantages include its role to clarify the assessment criteria, reduces the marking load on the teacher and encourages intrinsic rather than extrinsic motivation (Institute for interactive media and learning, 2006).

### 5.2.3.2. Implementing self and peer assessment

To involve students successfully in their own assessment, it is important to:

1. brief students and fellow teachers thoroughly before introducing the processes, making it quite clear in advance what is expected of them;

2. explain carefully the purposes of self and peer assessment to all parties, so that they do not see it as a dereliction of your duty as a tutor. Point out to students the benefits of becoming involved in assessment;

3. make sure that students are working with explicit criteria for success, either provided by teachers or negotiated with them;

4. ensure that whenever students are evaluating work they provide full and appropriate evidence for the marks or awards given, based upon the agreed criteria;

5. provide opportunities for rehearsal of the process in stress-free contexts. If possible, let there be a dry run which doesn't carry marks or at least begin on a small scale so that hiccups don't become disasters;

6. collaborate with colleagues who have already used self and peer assessment if possible, so that you don't try to reinvent the wheel and can learn from their mistakes;

   and have into consideration not to get everything right at the first time. Note what worked and did not work in the first instance and build in the results of the learning experience to the assignment with the next cohort of students. (Brown et al, 1994)

## 5.3.4. Journal writing

Journals are a students' record of a continual document of their expressions, experiences, feelings and reflections. Students can see their process and progression from reading their journal by the end of the assignment. They express themselves and write about their interests, goals, desires, worries etc. The teachers collect student journals on a daily or weekly basis, read them and return them with written comments, feedback or advice. Much of the information contained in journals can be obtained using different methods like interviews and questionnaires. Journals differ from these methods in that they basically under the control of the student. Students can write whatever they want and however they want. They may need some guidance and encouragement when they begin to use journals in order to become familiar with them. Min (2002)

### 5.3.4.1. Characteristics of Journals

Journal writing have a number of important benefits . They can be enjoyable, since they give the students free rein to write on any topic at the spur of the moment. In addition, they offer students the privacy, freedom, and safety ton experiment and develop as a writer (Cobbine, 1995 ).

Aschbacher, etal (1995) stated that journals are interactive in nature as they consist in written conversations between students and teachers. They provide a setting in which students not only display their knowledge, but reflect on the learning activity and ask questions or indicate a need for additional help.

Min (2002) mentioned other benefits like providing an aid to the memory since journal writers learn the value of recording their ideas for future use. They enhance critical thinking skills and provide support for self-directed learning activities.

Moreover, they provide psychological/emotional advantages by enabling individuals to work through difficult work or personal situations that can promote healing and growth.

Besides providing teachers with opportunities to assess their students and keep record of the learners' writing skills, dialogue journals give students opportunities to use language for genuine communication and enhance student involvement in and ownership of learning (Lafi, 2002).

Using journals allows the teachers to individualize language teaching by modeling writing skills in their response to students journals (Genesee and Upshur,1996: 120).

To ensure that journals are truly interactive and that they not become like other classroom assignments such as essays or reports, it is important that they not be highly structured or that limits  not be put on what students write about (ibid).

Teachers' comments about the learners' performance should not be judgmental. They should focus on meaningful communication and be supportive and only indirectly evaluative ( for example, " I am not sure what you mean by this; can you say it in another way?") (ibid).

### 5.3.4.2. Types of Journals

Lafi (2002) classified journals into two types, reflective and dialogue journals . In the reflective journals, students are encouraged to record their responses, identify their own values, attitudes and beliefs underlying their reactions to various learning situations and to reflect on how such values, beliefs and attitudes might affect their studies and their world in general. In dialogue journals, participants became involved in academic discourse and shared perspectives with their peers. Students are assigned

27

dialogue partners and in groups of threes were instructed to share with each other on course related issues. The instructor had full access to the ensuing conversations as well, rounding out the discourse community.

### 5.3.5. Interviews and Student-Teacher Conferences

Teacher-student interviews or conferences are productive means of assessing individual achievement and needs. They can include individual students, several students or the whole class. The most important about these types of assessment is the student and teacher work in collaboration in order to ascertain the progress the student has made.  In interviews, Students are given an opportunity to talk to the teacher in private.  Teachers  can identify what students have learned and what they have yet to learn. Conferences are an opportunity for students to present their work and to gain firsthand knowledge about how their work is assessed. Any disagreements between the student's self-assessment and his or her formal grade can be clarified.

It is important for students to feel comfortable with the interview process. Students should understand the objective of the interviews and should be aware that the teacher will be taking notes. Students will accept interviews as unobtrusive, routine classroom experiences if they are conducted regularly (Kulm, 1994 cited in NREL, 2004).

### 5.3.5.1. Characteristics of Interviews and Conferences

Shaaban (2001) & Assessment Companion (2002)  assured that interviews and conferences can be an effective informal way of assessing a student's progress in language learning. They provide opportunity for one-to-one interactions where the

teacher can learn about a student's communicative abilities, emotional and social well-being, attention span, attitudes, pace of learning, strengths and weaknesses etc.

The teacher can build on the information he collects from interviews and conferences to make decisions about instructional planning. Whether formal or informal, they provide students with personal attention and an avenue for two-way communication. Students can describe what is and is not working for them (Vose, 1997).

Interviews provide an opportunity for students to think at a high cognitive level. It is suggested that teachers begin a questioning sequence with general questions and proceed to increasingly specific questions. It is important to give students enough time to think through their answers, to reconsider them, and to respond to additional probes ( ibid).

Conferences provide important information for instruction because they set the course for future improvement and growth. Teachers can provide students with strategies that will improve their performance and help them to generate goals and next steps (Darling et al, 1995 cited in NREL,2004).

### 5.3.6. Portfolios

A portfolio is a planned collection of learner achievement that document what a student has accomplished in a given academic area( Borich, 2004). Students portfolio includes information about their strengths, needs, services provided, and support required and other related information about the learners and their  progress in mastering different writing skills. The students and their teachers look at the work collected in portfolios, to form a picture of what has developed, where difficulties remain, and direction for future work. (Allen, 1998:7).   In terms of writing

assessment, a portfolio is a collection of written texts for different purposes over a period of time. Normally, it takes one year. ( Weigle, 2002: 198)

A portfolio can be a file folder, a box or any durable and expandable multi-sourced container in which the student keeps:

- samples of writing

- lists of books

- book reports

- audio-recordings of speaking samples

- samples of tests

- self- assessment checklists

- samples of projects and artwork

- anything the student thoughtfully chooses to add to it Lafi (2002).

### 5.3.6.1. Characteristics of portfolios

Some Characteristics of portfolios assessment are:

1. Provides structure for involving students in developing standards for quality performance

2. Improves students' metacognitive ability to understand their own learning processes

3. Promotes integration of various learning activities and assessments

4. Enhances awareness of strategies for thinking and producing work

5. Creates documentation to submit to authentic audiences and/or reviewers to trace a student's progress over time.

6. Increases student accountability for their own learning

7. Promotes assessment of a wider range of learning styles

8. Encourages students' active involvement in the assessment process 12

9. Promotes self-assessment

10. Promotes authentic assessment of valued knowledge and skills (Park University,2003).

## 5.3.6.2. Types of portfolios

The Electronic Learning Community divided  portfolios into three different types:

**5.3.6.2.1. Documentation Portfolio**: This approach involves a collection of work over time showing growth and improvement reflecting students' learning of identified outcomes. It can include everything from brainstorming activities to drafts to finished products.

**5.3.6.2.2. Process Portfolio**: This approach documents all facets or phases of the learning process. It can show how students integrate specific knowledge or skills and progress towards both basic and advanced mastery. Additionally, it emphasizes students' reflection upon their learning process, including the use of reflective journals, think logs, and related forms of metacognitive processing.

**5.3.6.2.3. Showcase Portfolio**: This type of portfolio is best used for summative evaluation of students' mastery of key curriculum outcomes. It should include students' very best work, determined through a combination of student and teacher selection. This type of portfolio may include photographs, videotapes, and electronic records of students' completed work ( PGCPS, 2006).

### 5.3.7. Checklists

The are lists with columns for marking yes and no. They are often used for observing performance in order to keep track of a student's progress or work over time. They can also be used to determine whether students have met established criteria on a task.

To construct a checklist, Teachers need to identify the different parts of a specific communication task and any other requirements associated with it. "They consist of predesigned categories for recording observation .Thus ,they require precise and well-articulated categories and criteria for observing and assessing student performance or instructional activities  (Genesee and Upshur , 1996: 87).

### 5.3.7.1. Types of checklists

In some cases, a teacher will use a checklist to observe the students. In other cases, students use checklists to ensure that they have completed all of the steps and considered all of the possibilities.

### 5.3.7.2 Characteristics of checklists

According to NCLRC (2004) checklists can be useful for classroom assessment because they are easy to construct and use, and they align closely with tasks. At the same time, they are limited in that they do not provide an assessment of the relative quality of a student's performance on a particular task.

**Table(3)**
**Checklist to assess a student's handwriting by: Gomez etal (2005)**

| | |
|---|---|
| Student Name:……………………………… | |
| Date or Time Period of Assessment:………… | |
| Description of Writing Sample:……………… | |
| Check (X) appropriate criteria………………… | |
| | Problems with spacing between sentences, words, or letters |
| | Problems keeping 'on the line' |
| | Evidence of large amount of erasing or scribbling out words |
| | Lower case letters evenly made |
| | Incorrect formation of capitals |
| | Capital letters correctly formed |
| | Writing consistent in cursive letters |
| | Writing shows a mixture of cursive and manuscript letters |
| | Slant of the letters consistent within a sample of writing |
| | Slant of the letters inconsistent within a sample of writing |
| | Certain letters consistently malformed |
| | Able to produce a reasonable amount of writing in the time period   allotted |

### 5.3.8. Anecdotal Records

An anecdote is a short account of an incident. They are a teachers notes including specific dates, times and events of incidents that occur throughout the school day. Paulson (2002) states that anecdotal record is a collection of written observations of students related to their progress in learning . It may be kept in a separate notebook or included in a student's portfolio. These descriptions of student's activities and/or behaviors are done briefly and informally using only key words relating to the observed incident.

Teacher notes to students, whether offering criticism or encouragement, and student notes to teachers should also be part of the anecdotal records, as well as teacher annotations on a student paper . As anecdotes are complied over time, a teacher may be able discern patterns developing in a student's behavior and/or learning. Using this technique allows teachers the opportunity to modify their instruction to better meet the needs of their students (Worley, 2001).

### 5.3.8.1. Types of Anecdotal Records

Teachers use different techniques to execute anecdotal records. They may choose to keep running written observations for each student or they may use a more structured approach, constructing charts that focus each observation on the collection of specific data. A combination of open-ended notes and structured forms may also be used. All recorded observations are usually dated.

### 5.3.8.2. Characteristics of Anecdotal Records

Batstone (2004) affirms that a fundamental purpose of assessment is to communicate what the learner knows and is able to do. Teacher-generated, anecdotal records provide an insider's perspective of the learner's educational experience. This perspective is vital to communication with the learner and the learner's family. Anecdotal records also facilitate assessment conversations . Following is An example of an anecdotal record as presented by Batstone (2004)

**Table (4)**
**An example of an anecdotal record by Batstone (2004)**

| Name: _____ | Name: _____ |
|---|---|
| Date of assessment: _____ <br> Observation period: _____ <br> Comments:_____ | Date of assessment: _____ <br> Observation period: _____ <br> Comments:_____ |

### 5.3.9. Rubrics

Rubric is a assessment tool used to measure students' work, usually handed out before the assignment begins in order to get students to think about the criteria on which their work will be judged. It provides the learner with a clear picture of their learning and of areas for potential growth (Kennesaw State University,2006).

Andrade ( 2001) considers a rubric as a scoring tool that lists the criteria for a piece of work, or "what counts" (for example, purpose, organization, details, voice, and mechanics are often what count in a piece of writing); it also describes levels of quality from excellent to poor.

### 5.3.9.1. Types of Rubrics

There are two types of rubrics used to assess students' writings:  the holistic rubric and the analytic rubric.

In the holistic rubrics, student's work are scored as a whole, without judging the individual criteria separately. The aim is to provide the overall impression of a student's performance on a task.. Deficiencies in some aspects are tolerated as long as overall quality is high.

In the analytic rubrics, teachers score the individual criteria separately, then combines scores to obtain an overall total. It consists of multiple, separate scales.  The multiple scales enable students to pinpoint their strengths and weaknesses related to each criterion.

### 5.3.9.2. Characteristics of Rubrics

According to Min (2002)  rubrics have the potential to be excellent assessment tools because they offer students a vision of what the teacher is seeking to accomplish in the class and why it is important. The use of rubrics is one way to help promote effective evaluation procedures that reduces subjective grading procedures and offer student relevant information on their academic performance.

Allen (2004) points out that rubrics are explicit schemes for classifying products or behaviors into categories that vary along a continuum. They can be used to classify virtually any product or behavior, such as essays, research reports, oral presentations, and group activities. A rubric provides an instrument for student feedback that promotes assessment of learning. Table (5) which is adapted from Firewise Communities (2007) presents an example of a rubric used for assessing a writing task.

**Table (5)**

**An example of a rubric used to assess written a paragraph adapted from**

**Firewise Communities (2007)**

| CATEGORY | EXEMPLAR | ACCOMPLISHED | DEVELOPING | BEGINNING |
|---|---|---|---|---|
| **Content** | Includes all essential content. Ample evidence of analytical thinking skills. | Includes most essential content. Evidence of analytical thinking skills. | Includes some essential content. Some evidence of analytical thinking skills | Does not include sufficient content No evidence of analytical thinking skills. |
| **Topic Sentence** | Topic sentence is strong, clear, correctly placed, and restated in the closing sentence. | Topic sentence is clear, correctly placed, and restated in the closing sentence. | Topic sentence is somewhat unclear, incorrectly placed, or not restated in the closing sentence | Topic sentence is unclear, incorrectly placed, and not restated in the closing sentence. |
| **Supporting Details** | Paragraph has three or more supporting detail sentences that relate back to the topic sentence. | Paragraph has one supporting detail sentence that relates back to the topic sentence. | Paragraph has one supporting detail sentence that relates back to the topic sentence. | Paragraph has no supporting detail sentences that relate back to the topic sentence. |
| **Writing Mechanics** | All grammar, spelling, punctuation and capitalization are correct. No errors in text | Includes 3-4 grammatical errors, misspellings, punctuation or capitalization errors. | Includes 3-4 grammatical errors, misspellings, punctuation or capitalization errors. | Includes 5 or more grammatical errors, misspellings, punctuation or capitalization errors. |

According to the table student performance can be rated on a continuum from exceptional to insufficient. students can understand specifically what they are to do.

## 6. Conclusion

Tests are just one type of assessment, although they often get the most publicity .Many things students learn cannot be tested with a paper-and-pencil test. Lots of our students suffer from test anxiety which might lead them to failure on tests or cheating in order not to fail.

Proponent of alternative  assessment methods may claim that they require large amounts of time to develop and integrate into the curriculum. They also cost more because they require more sophisticated teaching, staff development, and scoring. They take not only time but energy to re-educate teachers, parents, and students in new forms of assessments.

Alternative assessment, however, holds great promise for English teaching and learning. Although the challenge to change existing methods of assessment and to develop new approaches is not an easy one, the benefits for both teachers and students are great. Students are given a diversity of learning opportunities to display knowledge, critical thinking skills, connect learning to their daily lives, and foster both individual and group oriented learning activities. If instruction focuses on the test, students will not learn these skills, which are needed for success in school and in life.

When adopting these alternative methods of assessments, teachers are offered unique opportunities to create relevant work that promotes academic achievement and individualizes the educational process. Moreover, they are provided with new perspectives on student learning such as insights to their individual learning styles.

According to General Administration of curricula (1999:69) writing is viewed as a process rather than a product therefore; traditional tests cannot assess different steps included in this process such as pre-writing, drafting and revising. This confirms the

need for using alternative assessment methods to show students progress in these different stages.

## Section II

## Previous Studies

### 1. Introduction

This section reviews 34 previous studies that former researchers have done about alternative methods of assessment. Brief details are given and both suggestions and recommendations of their studies are presented through the discussion. Then the researcher comments on these studies which are divided into three thematic categories.

The first type comprises studies about assessment in general. It presents some new methods used in assessing students learning in general. It discusses some advantages of these methods and some constraints towards their application. In addition, it provides an understanding of current assessment practices.

The second type includes 8 studies deal with language assessment. It probes some beliefs about language assessment, effects of using these methods on language instruction and how we can link assessment and instruction for better teaching and learning of our students.

The third type includes more than half of the studies which all focus on the assessment of students' writings. It explores the role of some assessment methods for developing students' writing ability. Further, it discusses the effectiveness of some programs designed to assess and develop students' writing.

## 2. Studies related to assessment

The section includes 8 recent studies that deal with alternative methods to assess students in different subjects other than English language. Some explore the possibility of adapting alternative assessment methods to assess students  and the advantages they possess. Others investigate teachers' perceptions about the best methods for assessing students according to their experiences,  in addition to students' perceptions of implementing given methods of assessment. The studies probe the general tendency to substitute the traditional methods of assessment with more authentic and effective ones. They  are sequenced chronologically in ascending order.

Plimmer (2000) sought to investigate some advantages of implementation of portfolios in a Computer Programming Course, the students' reactions towards this type of assessment and how well the portfolios meet the learning requirements of the courses. The sample of this study was twenty two students at Manukau Institute of Technology. In the first course, the students were required to submit their portfolio three times during the semester. For each submission they included pieces of work that demonstrated mastery of specific topics. The portfolio was 60% of the course mark. There was also a final examination that contributed the other 40%.

In the second course, the portfolio was used for the assessment of the learning of Visual Basic. This constituted 25% of the final mark. The students in this course were 2nd and 3rd year students who had done programming, systems analysis and database courses (as a minimum). The results showed that this portfolio was very successful at getting the students competent with Visual Basic. The average students enjoyed it and found it an easy way to learn

The researcher concluded that the advantages of portfolios were that they encouraged the students to actively review and select their work. This review process engages them in the higher order cognitive activities. However, he expressed his belief that portfolios are most appropriate when used in conjunction with other assessment methods such as tests or examinations.

Connor et al (2001) investigated the impact of brief daily essay quizzes as a strategy for assessing and promoting student learning. The study investigated whether students who took daily essay quizzes demonstrated better reading and thinking skills than students who took scheduled tests.

Participants were students in four upper-level undergraduate psychology classes. Two of the classes had taken four scheduled tests (ST) over the course of the semester The third class took weekly essay quizzes beginning halfway through the semester(a total of seven essay quizzes) in addition to the scheduled tests (ST/7EQ). The fourth class took essay quizzes every class day throughout the semester (DEQ). Each quiz consisted of one or two questions that tapped several levels of Bloom's taxonomy but could be answered and graded quickly .Quiz grades were based on content and clarity of ideas, not grammar or spelling. Students took the quiz at the first 5-10 minutes of each class. After turning in their responses, the quiz questions served as the means for beginning class discussion.

At the end of the semester, all participants read the same research article. The article addressed a topic that was not covered in any of the four classes, and none of the students had previously read the article. Copies of the article were distributed to students and they were asked to carefully read it in preparation for an essay quiz two days later. An essay question was constructed and distributed among the students.

An analysis of variance (ANOVA) indicated a significant difference in essay scores among classes. The scores for the DEQ class were significantly higher than those of both of the ST classes and significantly higher than the scores for the ST/7EQ class. There were no differences among scores in the two ST classes or the ST/7EQ classes

This study made clear that students who took daily essay quizzes showed better retention of information, clarity of ideas, and critical thinking when asked to write about an unfamiliar daily graded writing. These findings suggest that students who wrote daily essay quizzes went beyond simply learning the course material to develop reading and thinking skills that generalized to a new assignment outside their class. The daily quiz class had the opportunity to learn to read more carefully and critically because they were tested every class period. The study indicated that tests have a powerful role in shaping student behavior and skills. Accordingly, the researchers stressed the need for further examination of the role of testing in assessing and promoting student learning and thinking.

Wangchuk (2002) aimed at exploring the feasibility of adapting alternative assessment tools; namely rubrics and portfolio assessment for use in continuous assessment in Bhutan schools. Continuous assessment is defined in the study as a procedure which is concurrent with the course. It takes place at regular intervals throughout the course because it is impossible to assess al1 pupils all the time. Data were gathered using a questionnaire and focus group interviews. Two groups included in the study .A group of seven teachers and another group of five administrators. They were interviewed after the tabulation of the questionnaire responses of the participants.

Despite the identification of some difficulties, there was a broad consensus that both rubrics and portfolio assessment could be feasible and applicable in Bhutanese schools.

The two groups identified certain conditions for the successful implementation of rubrics and portfolio assessment. These conditions are: a) human resources could be developed   b) teachers could work collaboratively and spare some of their free time for rubrics and portfolio assessment c) teachers willingness to practice development and use of rubrics and portfolio assessment  d) materials should be provided to schools e) teachers should be trained and f) duties should be properly given to teachers. The researcher concluded that in due time rubrics and portfolio assessment can become very effective continuous assessment tools for teachers in the Bhutanese education system and promise solid pedagogical results in the long run.

Patterson and Bellaby (2002) aimed at examining student perceptions of Computer-Aided Assessment (CAA) in the Level One Systems Analysis module at the University of Derby. Primary research was completed principally by way of questionnaire, informal discussion with participating students, and observation of process and participants. Test results showed that many of the commonly perceived advantages of CAA, such as immediacy of feedback and objectivity of marking, were highly valued by students. 81% of students attached a high level of importance to immediacy of feedback, with 44% selecting the highest position on the scale of importance. A significant percentage of students (87%) thought that this type of assessment was "fair". There was a similar response (88%) when asked whether it was felt that CAA is a valuable aid to learning. In discussions with students it became clear that students take for granted a lack of bias. The results showed that there is a

certain degree of objectivity assumed by students. This makes CAA capable of fulfilling this expectation more adequately than might be the case in other forms of assessment. Student felt that they made some investment in their grade and that it has also caused them to prepare and review the work covered to that point In general, students in the survey were very positive about CAA.

Susuwele's study (2005) aimed at investigating teachers' perceptions of classroom assessment in mathematics and their current classroom assessments practices. The study sought to gain an understanding of the extent to which teachers use different classroom assessment methods and tools to understand and to support both the learning and teaching processes.

The study used a questionnaire, a lesson observation protocol, and pre-lesson and post-lesson observation interview protocols as main sources of data collection. The data collected through observations and interviews helped to elaborate the difference between perceptions of classroom assessment and the teachers' classroom assessment practices. Document analysis was used to provide information on the kind of written feedback students get and the nature of activities they do. A total of six teachers (three male and three female) from two primary schools in Malawi participated in this study.

The data showed that teachers perceive classroom assessment as tests that teachers give to their students at specified time intervals. In most cases, teachers assessed in order to rank students and not to identify individual capabilities and weaknesses. Five of the six teachers perceived assessment as testing, and classroom assessment practices were not clearly exist in their teaching. It was clear through the study that what teachers said about their teaching was not reflected during their

teaching. They showed limited ability to use different methods and tools to assess their students while teaching. There was very little attempt to understand how the students were learning.

Heidi & Du (2005) showed how using rubrics to assess students can support their own learning and academic performance. The participants were fourteen undergraduate teacher education students (six female and eight male Caucasian, middle class Midwesterners) participated in topical interviews in groups. Each student had completed an educational psychology course with Dr. Heidi prior to the interviews. The course and field experience involved regular use of rubrics, including co-creating rubrics in class, formal rubric-referenced self-assessment, and teacher feedback. Both the male and female students talked at length about how they used rubrics and about the results of rubric use.

Students' comments regarding rubric use were consistently positive. They said that using rubrics helped them focus their efforts, produce work of higher quality, earn a better grade, and feel less anxious about an assignment. Students also noted that rubrics help identify strengths and weaknesses in their work when used to give feedback, and that knowing "what counts" made grades seem fair. Their comments also revealed that most of the students tend not to read a rubric in its entirety, and that some may perceive a rubric as a tool for satisfying a particular teacher's demands rather than as a representation of the criteria and standards of a discipline.

Noonan & Duncan (2005) conducted a study to explore the nature and frequency of high school teachers' use of peer and self-assessment. Data for this study were collected from a survey of 118 high school teachers' assessment practices in a mid-

sized urban school jurisdiction in Western Canada. The teachers teach different subjects. The survey included several open-ended questions on teachers' assessment practices of the 118 teacher responses. 110 were used in the data analysis. In addition to 34 forced choice items on grading and assessment practices, the survey included several open-ended questions on teachers' assessment practices.

The results of the analysis showed that 26 teachers (24%) reported they did not use peer or self-assessment. Of the remaining eighty four responses, fifty-four teachers (49%) reported using peer or self-assessment 'a little'. The third category included 30 teachers (27%) who reported using peer or self-assessment 'somewhat'. An analysis was also made to compare the use of peer and self-assessment in the academic areas (Mathematics, Science, English, Social Studies) and showed that Social Studies and English teachers used peer and self-assessment somewhat more frequently than the other teachers. Teacher comments on the benefits of using peer and self-assessment were in three broad themes. First, peer and self-assessment encouraged and facilitated student reflection on their achievement. Second, peer and self-assessment were useful in assessing group work and third, they were useful in assessing students.

Yates (2005) conducted a study to determine the most appropriate methods of assessment for online courses. The population included 371 online instructors who had taught an Internet course or a web-enhanced course during the 2004-2005 academic year at the 15 Western North Carolina community colleges. The instructors were surveyed and asked which methods of assessment they use in the online courses that they teach and how effective they perceive those methods to be in determining if the learning objectives have been met for the courses that they teach online. The

findings of this study indicated that there is a difference between some academic disciplines in relation to the type of assessment methods being used in online courses.

There is a difference in perceived effectiveness of assessment methods among the individual instructors surveyed. The most effective means of assessment as determined by the survey results is individual projects. The least effective method of assessment as determined by the survey results is self-assessment. The study's results clarified that several instructors still use objectively scored testing. They didn't consider it the most effective method of assessment but this happened because of time constraints. Other results confirmed that a variety of assessment methods need to be used within each Internet course to determine the effectiveness of the course. Besides, there was no difference in the assessment methods being used by those instructors who received training and those who did not.

This conclusion could be because the training received by most online instructors was in Blackboard and/or technology and not assessment methods. The survey data indicated that there was not a difference in assessment methods being used by instructors who have taught for more than three years as compared to instructors who have taught three or fewer years. However, there is a difference in some of the types of assessments being used by instructors who teach more than one Internet course per year. Furthermore, instructors who have a large number of students and/or course sections resort to objectively scored testing methods only because they do not have time to grade alternative assessment formats.

# 3. Studies related to language assessment

This section includes 8 recent studies that investigate new methods used in assessing language. They include attitudes and beliefs about language assessment, effectiveness of some assessment method on language instruction and how alternative methods can link assessment and instruction to improve teaching and learning of students.

Fall et al (1997) explored the effect of group discussion and large-scale language assessment programs on students' comprehension. Large-scale assessment programs are beginning to design group assessment tasks in which small groups of students collaborate to solve problems or complete projects. Approximately 5,000 10th grade students from Connecticut public high schools participated in a pilot of a 90-minute language arts test. A random sample of 300 responses for each test form was scored holistically.

The study compared student performance on language arts tests in which they either were or were not permitted to discuss the story they were required to read and interpret. The analysis compared the quality of student responses on test forms using group discussion and without using it. Further, it examined qualitative changes in students' responses before and after collaboration, and examined students' reflections about the impact of collaboration on their understanding of the story. The results showed that a 10-minute discussion of the story in three-person groups had a substantial impact on student performance and clearly indicated the importance of collaboration to improve students' performance in assessments.

Brooks (1999) conducted a study to investigate the attitudes of adult students in an ESL program towards performance-based assessment. The types of performance-

based assessment used in the program are portfolios, presentations, and participation. The participants in this study were (N=127) adult English as a Second Language (ESL) students. The three types of assessment (portfolios, presentations, and participation) versus more traditional types of tests were surveyed by means of a questionnaire and semi-structured interviews. The results suggest that the participants in the study perceived all four types of assessment positively. Of the three types of performance-based assessment, presentations received the highest number of positive responses. Students assure that portfolios and presentations gave students the opportunity to demonstrate their ability. Analysis of background variables suggests that there were interaction effects for level of language proficiency and home country with regard to attitude toward assessment type. Other biographical variables showed little or no relationship to attitudes. The researcher considered performance-based assessment to have high face validity in the program since students perceived the assessment process as a tool for learning.

Bacha (2001) used a survey to find out if there is and a difference between their grade expectations and the actual grades they earned within the current assessment procedures used for assessing their essays. 150 Freshman English students at the Lebanese American University were surveyed on their grade expectations. The survey was given two weeks before the end of the semester. They were requested to indicate the grade range they expected on course essays. Random interviews with students were carried out. Results showed that students' expectations are significantly higher than their actual proficiency levels. The researcher assured the need to develop valid and reliable assessment procedures and the need to raise students' awareness of their abilities. She added that teachers need to help students increase their awareness and

understanding of the proficiency levels required in writing essays. One way teachers can do this is by showing their students sample essays. In addition, teachers need to clarify criteria for the different proficiency levels for the various types of writing tasks assigned throughout a semester.

Chen (2002) sought to investigate how English teachers in Taiwan junior high schools perceived the impact of a public examination, called the Basic Competency Test (BCT), on their curricular planning and instruction. The phenomenon of how external tests influence teaching and learning is described as "washback" in language instruction. The target population was Taiwan junior high school English teachers. The survey method (a quantitative method) and focus group interviews (a qualitative method) were used to collect data. 11 schools were selected and 151 teachers were asked to respond to the questionnaire .The interviewed teachers were selected from Taiwan junior high schools. Three groups of teachers, with five participants in the first group, six in the second group, and five in the third groups, were interviewed.

Findings from this study indicate that the BCT has an influential impact on teachers' curricular planning and instruction. However, such a wash back influence on teachers' teaching attitudes is quite superficial; that is, the washback may influence teachers what to teach but not how to teach. The researcher attributed that to the lack of in-service teacher training, teachers lack knowledge of how to change their teaching methods in order to align with the new curriculum. Based upon the findings, this study recommended: 1) provide teachers with extensive professional development opportunities, 2) change Teacher Monitoring and Evaluation Policy, 3) practice mix-ability grouping instead of achievement grouping to group students, and4) integrate assessment into classroom evaluation.

Abdel-Wahab (2002)  intended to examine the reactions of EFL students, EFL teachers, and school administrators to the use of unfamiliar assessment methodology , the self- assessment portfolio. The study employed  exploratory qualitative case study to examines the introduction of the self-assessment portfolio as a method of assessment in English as a Foreign Language (EFL) classes at Manarat Al-Sharqiah Intermediate School in the Eastern Province of Saudi Arabia. Data included the following: (1) EFL students' self-assessment portfolios; (2) interviews with EFL students; (3) interviews with EFL teachers; (4) interviews with Saudi school administrators; and (5) the researcher's observations and field notes. Eighty-one students were involved in the study. Forty-three students submitted self assessment portfolios. Data analysis revealed that most students enjoyed using the self-assessment portfolio and found the process of reflecting on one's own learning to be helpful. Moreover, a number of students who did not submit portfolios indicated in interviews that they thought this self-assessment approach would be helpful in identifying strengths and weakness in their learning. Interviews of two EFL teachers and three school administrators revealed that they consider the self-assessment portfolio as a type of assessment worthy of future consideration. Both portfolio and interview data suggested that the self-assessment method encourages students to adopt patterns of critical thinking and motivates students to learn.

Sook (2003) aimed at identifying the types of speaking assessment tasks used by Korean Junior Secondary School English teachers and the ways in which those assessments were administered. In addition, he aimed to investigate Korean teachers' perceptions of the practical constraints in Korean EFL classrooms which affect

assessment of speaking. The research was carried out using a questionnaire and interview method. A questionnaire was administered to ten English teachers who were working at ten Busan Junior Secondary Schools (Years 7-9) in Korea. Two males and eight females responded to the questionnaire. Four teachers who were representative of the ten original participants were selected for interview on the basis of maximum variation in age, gender, teaching experience, teaching setting, and grades taught.

The study revealed that Almost all teachers were using speaking assessment tasks which did not reflect authentic interaction between themselves and their students. Some constraints which affect the use of authentic interaction included: large classes, excessive work, face-to-face classroom teaching, lack of training in conducting speaking assessment and lack of effective and efficient instruments. Most teachers expressed a strong desire to learn how speaking assessment can be effectively and efficiently administered in the Korean EFL classroom context. The study recommended carrying out in-service teacher education programs, in which teachers have opportunities to retrain and refresh themselves in speaking assessment.

Heinz (2004) intended to determine if there is a significant correlation between recall protocol scores that are manually assessed and computer-generated recall assessment software package. The recall protocol demands that the reader comprehend the text well enough to be able to recall it in a coherent and logical manner. This procedure allows misunderstandings and gaps in comprehension to surface. The researcher developed a computerized recall protocol assessment software package as an extension and enhancement for the efficiency, consistency, and validity of an alternative measure.

The assessment measure was administered to 240 students studying German at the United States Air Force Academy in Colorado Springs, but 100 were randomly selected for inclusion in the actual study. Subjects worked independently on individual computers located in a central testing facility and the three original L2 texts were presented to the students on the computer screen in a random order. Once they indicated that they were ready to write their recalls, the text disappeared and they were prompted to enter their recall into the computer in their native English. Once parsing and spell checking were completed, the recalls were submitted to the scoring program and automatically scored. Concurrently, the recalls were manually scored in a traditional manner. The automated and manually scored recalls were submitted to correlational analysis and the automated recalls were further submitted to item and qualitative analyses. The study showed similarities of the results between the two procedures. Thus, both scoring systems provided valid and quantifiable data that generate similar information on the ability level of the subjects. Findings showed that the computerized procedure provides efficiency in delivery and scoring, enhances consistency, is practical for large-scale assessment, and can lead to improved diagnostic and placement testing. One advantage of using the automated system to generate scores clearly became clear when scoring times were examined. Manual scoring required an average of 2.8 minutes per recall while the automated recall was scored in an average time of 32 seconds per recall. Other advantages presented in its potential to reduce scoring subjectivity and to provide a window into the L2 reading comprehension process.

Yueming (2006) examined English as a second language (ESL) teachers' perceptions of classroom-based reading assessments. , their understanding about the function and

effectiveness of this assessment method, and the factors influencing this assessment process are presented. Six middle school and 7 elementary school ESL teachers participated in this study. Data consisted of interviews with ESL teachers, classroom observations, and assessment materials teachers used in the classrooms. Findings of this study include that ESL teachers highly value classroom-based reading assessments, considered them accurate and valuable and thought these assessments could provide great help to the daily teaching of reading. Teachers viewed state-mandated standardized testing negatively and of little value for English language learners.

## 4. Studies related to assessment of writing

This section includes 18 recent studies which focus on the assessment of students' writings. They explore the role of some assessment methods for developing students' writing ability. Further, they discuss the effectiveness of some programs designed to assess and develop students' writing.

Gearhart et al (1993) investigated the feasibility of using portfolio assessment as a method of evaluating elementary students' competence in writing. The study took place in primary school classrooms in Westerville, Ohio. The study contained two components: (a) an empirical study of the utility and meaningfulness of using a holistic/analytic rubric (developed for evaluation of traditional writing samples) to score students' portfolios; and (b) a qualitative analysis of scoring approaches, drawing particularly on raters' critiques of the analytic scoring approach. The portfolios are composed of both a "working" file and a smaller "showcase" file of students' selections of their best pieces. The three raters participated in this study were teachers experienced in using the analytic rubric for scoring district's assessments of students' narrative writing competence. Results show that the multiple samples contained within a portfolio provide a more comprehensive basis for judging writing quality and thereby support uniformity of judgment. However, raters sometimes rate collections higher than the average of their ratings of single pieces and this suggests that a collection may provide a context for anchoring judgments of the better pieces in the collection. Moreover, the design of a rubric must be coordinated with the design of a portfolio collection. Portfolios should be displays of work that teachers and students believe reveals students' competence along dimensions assessed by raters and known and understood by teachers.

Herman and Gearhart (1996) intended to illustrate techniques for establishing reliability and validity of assessments of students' narrative writing with holistic scales of two rubrics. A new rubric designed for classroom use and known to enhance teacher practice, and an established rubric for large scale writing assessment .The narrative samples were collected from elementary school located in the middle-class suburb in California. There were three data sets: direct assessment, sample of narratives which students wrote for classroom assignments, and narrative collections which included all of narratives written by each student with a range of 3 to 6 narratives. Results provided good evidence for the reliability and developmental validity of the new rubric, while correlations patterns were not clear. The researchers concluded that the holistic scale of the *Writing What You Read Narrative Rubric* can be used reliably and meaningfully in large-scale writing assessments of narrative collections and that can guide the work of teachers in classroom to produce performance –based assessments of writing that are both technically sound and usable.

Russell and Haney (1997) sought to compare between student performance on test conducted via computer and via paper-and-pencil. Interest in authentic assessment has increased within the educational community. To enhance the authenticity of tests of writing, as well as of other knowledge and skills, some assessments require students to respond in written form via paper-and-pencil. However, as increasing numbers of students grow accustomed to writing on computers, these assessments may yield underestimates of students' writing abilities. Two groups of students were randomly selected from grades 6, 7 and 8. For the experimental group, which performed two of three kinds of assessments on computer, 50 students were selected. The control group,

which performed all tests via pencil-and-paper, was composed of the 70 students. Findings show that, though multiple-choice test results do not differ much by mode of administration, for students accustomed to writing on computer, responses written on computer are substantially higher than those written by hand (effect size of 0.9 and relative success rates of 67% versus 30%). The researcher concluded that for students accustomed to writing on computer for only a year or two, such estimates of student writing abilities based on responses written by hand may be substantial underestimates of their abilities to write when using a computer.

Snford (1997) examined the processes in which students' writings are assessed and the underlying beliefs that teachers hold about students' assessment, as well as the dilemmas facing teachers and students relating to the purposes of evaluation. Twenty high school students in Alberta and five of their high school teachers participated in the study. The information in this study were gathered through observation, informal and formal interviews. Data analysis revealed that each teacher used a variety of writing assignments in their classes to assess their students' progress. Few teachers used alternative methods of assessment such as: journals, workshops, conferences and portfolios. The researcher stated that students still rely on teacher-assigned grades to determine their success. They cannot determine for themselves how well they have accomplished a task, because they have never had the opportunity. Teachers rarely modeled self-assessment in the classroom; therefore, students do not have a concept of how to talk about or to feel about assessing their own work. There is a need to allow students to express their views, make choices, take risks, and explore alternative styles. Assessment needs to become a major part of the learning process rather than the final act imposed on the products of Learning.

Andrade (1998) presented a study took place during the 1997-98 school year and looked at the effects of instructional rubrics and guided self-assessment on students' writing and understandings of good writing. This study involved thirteen seventh- and eighth-grade classes in the same two urban schools. Both the treatment and control groups wrote two essays: a historical fiction essay, and a response to literature. Students in all participating classes were given instructional rubrics, but only the treatment classes were engaged in a process of guided self-assessment. The two self-assessment lessons focused on a formal process of guided self-assessment designed in collaboration with the participating teachers. During class, students were asked to underline "time and place" in red on their rubrics, then underline the information they provided about the time and place of their story in red on their essay. This process was followed for all seven criteria on the rubrics. Control classes received copies of the rubrics but did not formally assess their own work in class. The results of study indicated that rubric-referenced self-assessment can have a positive effect on girls' writing but no effect on boys' writing. This study did not examine students' cognitive and emotional responses to self-assessment. The study recommends that there is a need for better understanding of the different ways in which boys and girls respond to self-assessment.

Curfman and Crehan (1999) investigated the effect of rapid feedback for a state writing assessment on the quality of students' writing and how closely the classroom teacher's score agree with the state department of education scores in Nevada. Eighth grade English teachers (n=8) were trained in analytic scoring of writing assessments. They then scored their own students' state writing assessments after administration of

the assessment. They also scored assessments for a partner teacher's class. A second writing assessment was administered to the classes of the eight participating teachers and to eight control classes. At about the same time a brief questionnaire was designed to assess teachers' attitudes towards the statewide assessment program. . Results show good agreement between the teachers' scores and the scores assigned by the state department. There was no difference between the writing performance for students of project teachers and students in the control classes. Teachers thought that the writing assessment was useful and they also expressed the desire to be trained more in the analytical scoring method and in the teaching writing skills. The researchers suggested to train a teacher at each school to be "scoring leader" to coordinate training and scoring of the assessment at the school site.

Dorji (2000) explored the way Bhutanese grade seven and eight teachers of English respond and use rubrics to modify their instruction and assessment practices. A group consisting of five teachers from two high schools and a junior high was used. The data for the study was generated through focus group discussions, individual interviews, classroom visits, as well as reflections on his own teaching experiences.

The researcher described the response of the teachers as very optimistic and enlightening. Their responses clearly indicated benefits derived from the use of rubrics. Rubrics made their instruction effective. They were able to make their expectations clear and to focus and guide students how to meet these expectations. Additionally, using rubrics allowed teachers to focus their instruction on students' specific needs through mini lessons and individual guidance and to monitor the progress students made. Teachers found rubrics to be an effective tool to guide self- and peer- assessment by increasing students' sense of responsibility for their own

work. The most important finding as expressed by the researcher was that rubrics paved the way for teachers to use the writing process approach and value the process rather than the product in student writing. Some recommendation included in the study emphasized that : a) teachers need to be further trained to develop and use rubrics, b) teachers need to be trained to teach writing through the writing process approach, and c) schools need to adopt a new outlook towards writing and assessment.

Mnuputty (2000) initiated a pilot study to investigate whether or how students become better at self assessing their own learning, and how students writing performance improved over a semester with the use of self-reflection . The participants in the study were six students selected from students enrolled in a Writing class at Pattimura University in Indonesia during the Spring 1999 term. The entire class received process writing instruction and produced a set of six writing tasks accompanied by six written reflections on learning, during a total learning time of 14 weeks or one semester. The results of the study showed an increase in students' ability to self-assess their learning problems and needs as the semester progressed. Students were also found to have home better in setting learning goals and planning ways to implement their goals over time. Their writing also developed, particularly the compositional aspects of organization and development.

Erdosy (2000) sought to investigate the processes which raters of ESL compositions follow in constructing scoring criteria to assess a corpus of compositions, and how these processes can be related to their personal and professional backgrounds. The researcher analyzed data generated through questionnaires and think-aloud protocols

by a project to investigate raters' decision-making behaviors. In the project, raters were asked to assess 60 compositions without relying on a rating scale. Subsequently, the influence of background factors on this process was explored through interviews with the participants, which were designed based on the analysis of the think-aloud protocols. Differences in scoring criteria emerged through differences in raters' backgrounds, particularly in their teaching experiences, which led them to differ in identifying the key competencies learners mastered at various levels of development. The most significant sources of variability in the case of the four participants in the present study lay in teaching experiences and, in the case of the non-native -speaking raters lay in learning experiences. Academic background did not emerge as a significant factor, and assessment experience was likewise limited in its impact, influencing mostly rating strategies, but not the establishment of scoring criteria.

Andrade (2001) examined the impact of instructional rubrics on eighth grade students' writing and on their knowledge of the qualities of effective writing. The research was conducted in nine eighth-grade classes in two different middle schools in Southern California. The sample from both schools included 242 students. Half of the students were boys and half were girls. They were asked to write three different essays approximately one month apart. Students in both the treatment and control classes were asked to write first and second drafts of the essays. Students in the treatment group were given instructional rubrics that articulated the criteria and gradations of quality for three assigned essays but the students in the control group were not given a rubric. The second tool used in this study was a questionnaire to uncover students' beliefs about "what counts" when evaluating an essay. Findings from the analysis of the essay scores indicated that rubrics can orient students toward the criteria for

writing as communicated by the rubric and can help students write to those criteria. Responses to questionnaires revealed that the students who received the three instructional rubrics had more (if not complete) knowledge of what counts in writing and of the criteria by which their essays were evaluated. The researcher concluded that instructional rubrics seem to have the potential to broaden students' conceptions of effective writing beyond mechanics to include qualities such as word choice, voice and tone.

Nakamura (2002) examined the strengths and weaknesses of holistic and analytic scoring methods and explores how holistic or analytic scales can be used to better assess student compositions. Ninety students took a composition test in class (30 students per class), and their writing scripts were evaluated by three raters both holistically (using one evaluation item = overall) and analytically (using five rating items chosen by the author: grammar, vocabulary, organization, originality, cohesion). The researcher stated that If large numbers of students need to be placed into writing courses with limited time and limited resources, a holistic scale may be the most appropriate choice in terms of practicality. But to avoid risky idiosyncratic ratings, analytic assessment (with several evaluation items) is strongly recommended. The more ratings a person receives, the higher the rating precision.

Reid (2003) aimed at exploring the role of Assessment Programme which was entitled 'Local Moderation' for encouraging teachers to use formative assessment to support pupil learning and for developing procedures to quality assure teachers summative assessment judgments. A team of primary school teachers and secondary school teachers of English in Scotland have developed skills in using criterion

referencing to assess the writing of pupils in the 5-14 age group, using the national testing criteria framework. The study sought to investigate how these skills can be further developed through assessment moderating sessions, meetings where teachers work cooperatively to refine and confirm assessment judgments. A group of four primary school teachers and two secondary school English teachers constituted the action research team. The researcher's role was to mentor and support the practitioners engaged in action research studies for the Local Moderation Project and investigating teachers' experience of participation in the project. Meeting notes and recordings were collected during the progress of the action research project Interviews were conducted at the start of the projects and towards the end as teams were writing case study reports. It was suggested that adopting such a model would help teachers resolve a perceived tension between summative and formative assessment demands in relation to the writing curriculum. The results show that, using assessment rubrics seemed to help teachers focus their teaching aims and perceive clearer links between those and assessment criteria. Negotiated rubrics increased the opportunities for peer and self assessment. Teachers felt that they were more in touch with learning while it was happening in the classroom. They seemed to feel more in touch with pupils as writers rather than simply assessing the product of pupils' writing. Pupils' metacognitive skills also improved. The rubrics enabled pupils to focus on specific aspects of writing while they were writing, thus acknowledging both the diagnostic and formative potential of assessment. Teachers' diagnostic assessment of writing was enhanced by making them commenting on specific improvements in aspects of writing, rather than in improved overall attainment: these improvements related to both compositional skills. Furthermore, the linear structure of the rubric seemed to encourage children to identify their own next steps rather than put a ceiling on achievement. The

relationship between learning intentions planned by the teacher and the success criteria by which their achievement would be judged, was clearer for children.

Chen & Wei (2003) investigated how Chinese learners of English at intermediate level were guided through the process of self assessment to develop their writings. The participants in this study were ninety students, fifteen female and 75 male. They were all postgraduate students at Shanghai Jiao Tong University. Four techniques were employed to help the students implement self assessment: checklists, teacher assessment, peer assessment and writing knowledge input. Two checklists were designed for the students to assist them focus on the overall aspects of composition and to help them focus on grammatical errors. The first piece of written work was collected and evaluated by the teacher then by classmates. Teacher assessment and peer assessment provided the students with assessment techniques they could transfer to their self-assessment. When the second writing task was done, students were required to submit their composition with a self- comment attached. The result of the study showed that the use of these techniques encourages students to look critically and analytically at their writing and to take more responsibility for what they write. It opened up opportunities for exchanges of opinions between teacher and student. In addition, it cost the teachers less time and reduced the teacher's workload.

The Student Learning Assessment Team of "Capital Community Collage "(2003) examined the effectiveness of an assessment  programme to develop students' writings. Twelve teachers in diverse fields gave a writing assignment to their students in fifteen different classes. Students in all classes received the same assignment sheet which asked them to read an article selected by their teacher and compose a response

to it. They were asked to devote their first paragraph to introducing and summarizing the information in the article and then to spend the rest of the paper presenting their thoughts about the topic. They were urged to support their ideas with references to the article, to other sources of information, or to their own experiences. In the project, every sampled essay was read twice, first holistically and then analytically, and then read twice again by a second reader. Essays were scored on a four-point scale, with levels identified as 1) in progress, 2) essential, 3) proficient, and 4) superior. The results indicated that the main obstacles to students' proficiency lie in the areas of development (support of ideas with evidence, examples, elaboration of topics, etc.) and language (effective use of sentence structures, word choices, and mechanics of standard written English).The study stressed the need of curricular planning and professional development activities that explore methods to increase students' skills in the categories of development and language and to be defined scoring rubrics. These rubrics should be distributed to students and staff throughout the college in order to open discussion of Capital Community College writing standards. In addition, the college should develop goals and practices for increasing the percentage of students writing and engage more faculties in assigning, assessing, and improving students' writing.

Srimavin and Darasawang (2004) aimed to investigate if journal writing enabled MA participants to develop self-assessment. The participants were one male and three females MA students at King Mongkut's University of Technology Thonburi. They undertook the course Self-assessment for four hours per week for 15 weeks in the second semester of the MA program. Each class was of two hours duration. The teaching in this course was discussion-oriented and based on the assigned reading and

experiences of the participants as learners and teachers. Journal writing was selected because it is a simple technique. The participants already knew the process of journal writing because in the first semester they had to do journal writing for the whole semester. After the second week of the course, the participants were asked to write a journal in English at the end of each class and to hand in their journals every week. They were asked to reflect on the course content. To help the participants reflect on the areas which activated their self-monitoring ability the tutor asked them some questions as a guide to writing their journals. Content analysis was used in the data analysis by looking at how the subjects responded to the questions given as guidelines. The findings show that depth in self-assessment did not develop. This might be because there was no reaction from the tutor/researcher to what was written in the journals and the subjects treated the journal as an exercise where they answered the guiding questions rather than a place to assess their own learning. The guided questions seemed to have obstructed the reflection process and the subjects did not realize the benefits of a journal as a tool for self-assessment. The researchers recommended that guided questions, should be provided at the initial stage of journal writing, if the learners are not used to reflecting on their learning or assessing their own learning. However, students should be asked to move away from the questions gradually or the teachers could change the questions. Furthermore, Negotiation between the teacher and the learners on what is to be assessed could help the learners perceive what is required.

Andrade and Du (2005) aimed at exploring the impact of rubrics on eighth grade students' writing and knowledge of the qualities of effective. The participants of the study were fourteen undergraduate teacher education students (six female and eight

male Caucasian, middle class Midwesterners) took part in topical interviews in focus groups. Three of the groups included four students. Because two women could not attend their scheduled focus group interview, one of the female groups had two students. Each student had completed Dr. Andrade's 200-level educational psychology course with a field placement prior to the interviews. The course and field experience involved regular use of rubrics, including co-creating rubrics in class, formal rubric-referenced self-assessment, and teacher feedback. The students stated that using rubrics helped them focus their efforts, produce work of higher quality, earn a better grade, and feel less anxious about an assignment. Their comments also revealed that most of the students tend not to read a rubric in its entirety, and that some may perceive of a rubric as a tool for satisfying a particular teacher's demands rather than as a representation of the criteria and standards of a discipline.

Gansle et al (2005) designed a survey to examine teachers' preferred assessment and scoring methods which best represent student writing abilities when using curriculum based measurement (CBM) and other assessment procedures. The researcher claimed that CBM had been validated for use as "dynamic indicators of basic skills". Teachers and classroom professionals are routinely taught to administer and score the products of CBM, and can do so reliably. It provides data from which the effectiveness of an intervention or instructional program can be determined. Thus it eases and standardizes the process of formative evaluation for improving students' academic outcomes

The sample in the survey consisted of 335 teachers from schools located in eight states in various geographic regions throughout the United States. Results showed that teachers perceive variables such as complete sentences, number of nouns or verbs,

and correct punctuation marks to represent student writing skill better than total words written, words spelled correctly, or correct word sequences. They believe that their holistic ratings of student writing skill outperform all other types of assessment.

Kenworthy (2006) examined the effects of additional time and different media on the overall quality of English language learner's written assessment tests. The participant were sixteen intermediate-level students (L1 Cantonese), enrolled at a campus of an American university within Asia .They were required to sit for a placement test to measure their overall English composition skills.  This test is a prerequisite for admittance into the Ohio University. First, they manually wrote a 45-minute timed placement test in the confines of an educational setting. Several weeks later these same students were allotted one week to complete a computer-generated essay at their personal residence. By contrast, there were statistically significant differences in the number of reported grammatical errors. On the whole, when compared to the timed writings, the at-home essays were characterized as having fewer numbers of grammatical errors and greater holistic scores which supports the idea that participants efficiently used additional time and electronic technology to affect overall quality of their texts.

# 5. Comments on the previous studies

The previous studies deal with different assessment issues in different concerns. They are conducted by different researchers in different countries at different educations levels: universities and schools. Similarities and differences emerged between the studies and also between them and this study.

The first part of the section included studies about assessment in general. It presented some new methods used in assessing students learning in different subjects. These studies revealed the movement towards adopting alternative methods of assessment to become a major part of the learning process rather than the final act imposed on the products of learning. The study of Connor et. al (2001) showed the impact of brief daily quizzes on undergraduate psychology classes while Heidi & Du (2005) explored the effect of rubrics on students who participated in educational psychology course. Some studies explained the use of these new methods in online courses like: Yates's study (2005) and the advantages of implementation of portfolios in a Computer Programming as in Plimmer's study (2000). Patterson and Bellaby's study (2002) examined student perceptions of Computer-Aided Assessment. Susuwele's study (2005) investigated teachers practice and perceptions of classroom assessment in mathematics.

The second part included studies about the assessment of language. It probed some beliefs about language assessment, effects of using these methods on language instruction. Some of the studies explored the effectiveness of certain programs in assessing language and their role for improving and developing student learning. Heinz's study (2004) discussed a program to assess reading comprehension using the computer, "A computerized Recall Protocol ". Chen's study (2002) investigated the

perception of Taiwanese teachers of the impact of a public examination, called the "Basic Competency Test" on their curricular planning and instruction.

Other studies dealt with the attitudes towards the adopting and implementing developed language assessment programs. They showed how assessment and instruction can be linked for better teaching and learning of students and how these methods motivated students and encouraged them to adopt patterns of critical thinking. The study of Brooks (1999) explored the attitudes of adult students in an ESL program towards performance-based assessment. Abdel-Wahab's study (2002) examined the reactions of EFL students and teachers, and school administrators to the use of the self- assessment portfolio. Sook's study ( 2002) investigated Korean teachers' perceptions of the practical constraints in Korean EFL classrooms which affect assessment of speaking .The study of Fall et al (1997) explored the effect of group discussion and large –scale language arts on students' comprehension.

The third part focused on the assessment of students' writings. It explored the role of some assessment methods for developing students' writing ability. The role of assessment programs for developing students' writings were discussed in some studies like the study which was conducted by the Student Learning Assessment Team of "Capital Community Collage "(2003). Reid's study (2003) explored the role of Assessment Program entitled 'Local Moderation' for encouraging teachers to use formative assessment. Russell and Haney's study (1997) compared between student performance on test conducted via computer and via paper-and-pencil. A study conducted by Gansle et.al (2005) examined teachers' preferred assessment and scoring methods which best represent student writing abilities when using curriculum based measurement (CBM) and other assessment procedures.

71

Other studies examined the impact of rubrics on students' writing such as : Andrade's study (2001) , Herman and Gerhart" study (1996) , Andrade and Du's (2005) and Dorji's study (2000).

The effects of some methods used in writing assessment were revealed in some studies. Curfman and Crehan's study (1999) investigated the effect of rapid feedback on the quality of students' writing. Srimavin and Darasawang'study (2004) investigated the role of journal writing for enabling MA participants to develop self-assessment. Mnuputty's study (2000) investigated the usefulness of criterion-reference assessment techniques for promoting students self assessment and their role to improve and progress their writing. Kenworthy's study ( 2006) examined the effects of additional time and different media have upon the quality of English language learner's written assessment tests. The study by Erdosy (2000) investigated the processes which raters of ESL compositions follow in constructing scoring criteria to assess a corpus of compositions.

Most of the previous studies made clear the need for assessing students depending on new methods of assessment to involve students in making decisions about the degree to which their performance matches their ability and not to rely on teacher-assigned grades to determine their success. This can be conducted using a variety of available instruments and methods as: journals, workshops, checklists, conferences                                   and                                   portfolios.

# Chapter III

# Methodology

1. Introduction

2. Research Design

3. Population of the Study

4. The Study Tools

5. Validity

6. Reliability

7. Procedures of the study

8. Conclusion

# Chapter III

## Methodology

### 1. Introduction

The purpose of the present descriptive analytic study is to identify the existing assessment tools used for assessing the 10$^{th}$ grade student's writing and the matchability of those methods with modern methods of writing assessment. Additionally, it probes some obstacles towards the adoption of the modern methods for assessing writing . To fulfill the study aims, the researcher used a questionnaire to collect the needed information.

This chapter discusses the population of the study, the steps  and the tools used to answer the study questions and to deal with statistical analysis, research methodology and procedure of the study.

### 2. Population of the Study

The population of the study consists of all the English language teachers of 10$^{th}$ grade at Gaza Government schools in (2006-2007). Table (5) shows the population of the study according to teachers' gender and directorate : (North Gaza , Gaza , Wostta , Khan-Younes  , Rafah )

**Table (6)**

**The Population of the Study**

| Directorate | Gender | | Total | |
|---|---|---|---|---|
| | Male | Female | Number | % |
| North Gaza | 22 | 19 | 41 | 16.14% |
| Gaza | 48 | 45 | 93 | 36.61% |
| Wostta | 23 | 19 | 42 | 16.54% |
| **Khan-younes** | 24 | 20 | 44 | 17.32% |
| Rafah | 20 | 14 | 34 | 13.39% |
| Total | 137 | 117 | 254 | 100.00% |

**2.1. The Sample of the Pilot Study**

The sample of the pilot study consists of (30) teachers of the study population. It was used to compute the validity and the reliability of the study tool. The piloting sample was excluded from the population when the main sample drawn.

**2.2. The Sample of the Main Study**

The sample of this study consists of (150) male and female teachers. The researcher received only (133) completed questionnaires from the teachers, (6) questionnaires were excluded because some data were missing. The final number of the teachers in the main sample was (127) teachers (approximately 52.4 % of the population ). The following table shows the sample of the study.

**Table (7)**

**The Sample of the Study**

| Directorate | Gender | | Total | |
|---|---|---|---|---|
| | Male | Female | Number | % |
| North Gaza | 12 | 12 | 24 | 18.9% |
| Gaza | 21 | 18 | 39 | 30.7% |
| Wostta | 12 | 8 | 20 | 15.7% |
| Khan-younes | 12 | 16 | 28 | 22.1% |
| Rafah | 8 | 8 | 16 | 12.6% |
| Total | 65 | 62 | 127 | 100% |

A graphical representation of the main sample of the study in the light of the gender and the directorate is shown in figure (1). This diagram shows that the highest number of the participants who filled in the questionnaire was in Gaza directorate, where 21 male participant and 18 female participants completed it. The least number was in Rafah. Only eight male and eight female participants did that.

Figure ( 1 )

The main sample in the light of Gender & Directorate
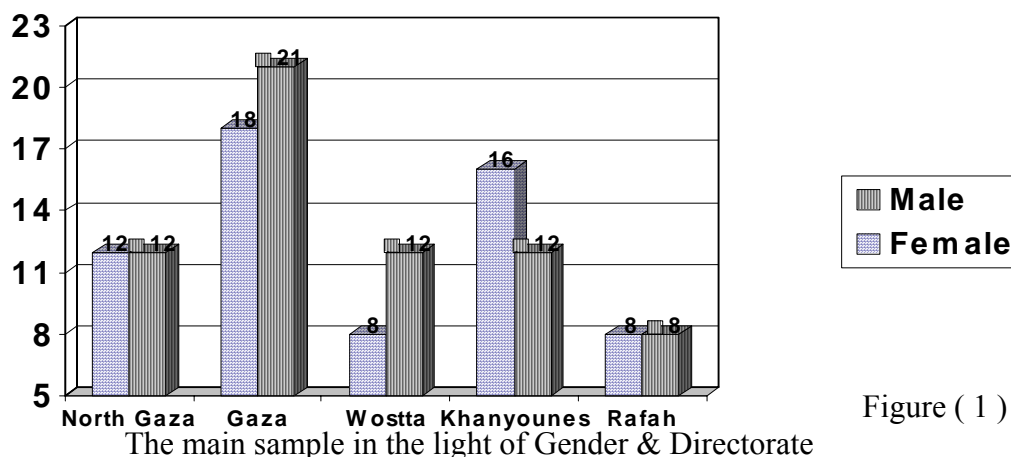
## 3. The Study Tool

The researcher used a questionnaire to fulfill the aims of the study and to answer its questions. The questionnaire was developed to be answered by the 10th grade English teachers in Gaza schools in the scholastic year (2006-2007).

## 3.1. The Questionnaire

The aim of this questionnaire is to help the researcher to evaluate the methods used in assessing the 10th grade students' writing with regard to the modern methods of writing assessment .

## 3.1.1. Questionnaire Construction

Based on the literature review and consultation with specialists, the researcher conducted a questionnaire which includes 10 domains of assessment that cover traditional and modern methods for assessing writing. Each of the ten methods contains different items which clarify the procedure by which the teachers apply this method of assessment (Appendix I).

75

## 4. Validity

Bynom (2001:1) defines validity as the truth of the test in relation to what it is supposed to evaluate. It is concerned with the relevance and usefulness of what you are measuring. The questionnaire was evaluated by a panel of specialists (curriculum designers, methodologists, supervisors and sixth grade teachers) from different institutes in the field of education .

## 4.1. Validity of the Questionnaire:

To examine the questionnaire validity, the following steps were used:

### 4.1.1. Content Validity (Experts Judgment)

The questionnaire was introduced to 10 EFL specialists, including university professors and highly qualified long experienced teachers of methodology, from different universities. Juries modified this questionnaire by adding, deleting or correcting its items. All juries approved the questionnaire, which the researcher applied it to his study.

### 4.1.2. Construct Validity:

### 4.1.2.1 Internal Consistency Method:

This type of validity indicates the correlation of each item degree with the total average of the questionnaire. The researcher used this method to calculate the correlation of the questionnaire. Pearson correlation formula was used to calculate the correlation between the score on each item of the questionnaire with total score of the questionnaire. Table (8) shows the correlation coefficient of the questionnaire.

**4.1.2.2 Internal Consistency Validity:**

An internal consistency validity was conducted through:

Firstly , the correlation (Pearson) between the items and their domains was calculated and it was found that all the items have a significant correlation coefficients at (0.05 & 0.01) levels with its domain. The correlation coefficients was between (0.421 – 0.703) which  is considered evidence for the consistency of the study tool.

Secondly, the correlation (Pearson) between the domains and whole tool was calculated. The following table shows the relation between the domains and the whole tool.

**Table ( 8)**
**The correlation coefficients between the Tool & its domains**

| The Domain | The Correlation Coefficients |
|---|---|
| Testing | 0.59** |
| Portfolio | 0.72** |
| Self assessment | 0.61** |
| Peer assessment | 0.78** |
| Journals | 0.69** |
| Interviews & Conferences | 0.81** |
| Rubrics | 0.67** |
| Observation | 0.73** |
| Anecdotal records | 0.80** |
| Checklists | 0.54** |

** The correlation coefficient significant at (0.01) level.

 The table above shows that all domains are significantly correlated with the whole tool, which means that the tool have internal consistency validity.

## 4.2. Discriminate Validity:

Discriminate validity is used to show if the tool can discriminate between the highest group (the highest 33% from piloting sample) and the lowest group (the lowest 33% from piloting sample). Mann-Whitney Test was used to compare between the two groups. The following table describes the results.

**Table (9)**
**Discriminate validity**

| Domain | Group | Mean Rank | (U) Value | (Z) Value | Sig. |
|---|---|---|---|---|---|
| Testing | High | 14.1 | 49 | 3.558 | 0.01 |
| | Low | 5.44 | | | |
| Portfolio | High | 15.5 | 55 | 3.916 | 0.01 |
| | Low | 5.5 | | | |
| Self assessment | High | 15.5 | 55 | 4.045 | 0.01 |
| | Low | 5.5 | | | |
| Peer assessment | High | 15.5 | 55 | 4.049 | 0.01 |
| | Low | 5.5 | | | |
| Journals | High | 15.5 | 55 | 4.058 | 0.01 |
| | Low | 5.5 | | | |
| Interviews & Conferences | High | 14.9 | 61 | 3.453 | 0.01 |
| | Low | 6.1 | | | |
| Rubrics | High | 12.5 | 85 | 2.163 | 0.05 |
| | Low | 8.5 | | | |
| Observation | High | 15.50 | 0.00 | 4.359 | 0.01 |
| | Low | 5.50 | | | |
| Anecdotal records | High | 15.50 | 0.00 | 4.359 | 0.01 |
| | Low | 5.50 | | | |
| Checklists | High | 15.50 | 0.00 | 4.359 | 0.01 |
| | Low | 5.50 | | | |
| Whole Tool | High | 15.50 | 55 | 3.8 | 0.01 |
| | Low | 5.50 | | | |

The table shows that all (Z) values are significant at (0.01 – 0.05) levels, i.e the questionnaire can discriminate between the two groups.

# 5. Reliability

Fraenkel and Wallen (1996:10) defined reliability as the degree to which scores obtained with an instrument are consistent measures of whatever the instrument is measuring. The researcher computed the reliability of the questionnaire through:

## 5.1. Cronbach Alpha Method:

"Cronbach Alpha" is considered the most general form of reliability estimates and it also, concerned with homogeneity of items compromising the scale" (Thorndike, 1997). The researcher used Alpha Cronbach Formulas to examine the reliability of this questionnaire.

**Table (10)**
**Cronbach Alpha Reliability Coefficients**

| The Domain | Alpha Coefficient |
|---|---|
| Testing | 0.79 |
| Portfolio | 0.84 |
| Self assessment | 0.85 |
| Peer assessment | 0.91 |
| Journals | 0.78 |
| Interviews & Conferences | 0.74 |
| Rubrics | 0.85 |
| Observation | 0.81 |
| Anecdotal records | 0.82 |
| Checklists | 0.84 |
| Whole Tool | 0.92 |

The results show that all the coefficient alphas had acceptably high internal consistency, given that the value of alpha should be at least 70% .This indicates that the questionnaire is highly reliable.

## 5.2. Split Half Method:

Table (11) describes Split Half Reliability Coefficients of the questionnaire which shows that the lowest level of scores was got by peer assessment (0.74) and the highest level of scores was got by the portfolio (0.92). These results confirm that all split half reliability coefficients are highly reliable.

**Table ( 11 )**
**Split Half Reliability Coefficients**

| The Domain | Split Half Coefficients |
|---|---|
| Testing | 0.85 |
| Portfolio | 0.92 |
| Self assessment | 0.91 |
| Peer assessment | 0.74 |
| Journals | 0.77 |
| Interviews & Conferences | 0.86 |
| Rubrics | 0.88 |
| Observation | 0.76 |
| Anecdotal records | 0.88 |
| Checklists | 0.79 |
| Whole Tool | 0.90 |

The results confirm that all split half rliability coefficient are more than 70% and highly reliable

## 6. Instructions of the Questionnaire:

The instructions of the questionnaire were clearly written in English on an attached paper. The instructions included the purpose of the questionnaire and guiding steps for the teachers to choose the appropriate answer for every item. (see appendix III )

## 7.  Procedures of the Study:

To fulfill the study, the researcher carried out the following steps :

- Collected and reviewed the previous related studies to get benefits from their procedures, tools, results, thoughts, and recommendations.

- Prepared the theoretical framework of the study through reading some specialized  books in this field and searching in previous related studies.

-  Prepared the tool of the study.

- Asked experts to review the questionnaire and then give approval to be applied.

- Took permissions From the Ministry of Higher Education to apply the study

- Applied the tool of the study on the $10^{th}$ grade English teachers in Gaza schools.

- Cooperated with school principals and the teachers to make sure that the questionnaire is answered .

- Collected the study tool and cooperated with a statistician to analyses data statistically for the result of questionnaire; using  SPSS  v. 11.

- Concluded the final results of the tool.

- Attempted to give explanations for the results and find some justifications and then give recommendations and suggestions.

## 8. Conclusion

In this chapter the researcher explored the population of the study, and the tool used to answer the questions of the study. A sample of (30) teachers was used as a pilot study to determine the validity and the reliability of the study tool.

Content Validity was approved by introducing the questionnaire to a panel of experts" 10 EFL specialists. Internal Consistency Validity the correlation (Pearson) between the items and their domains was computed, then the correlation between the domains and the whole tool. The results showed that the items are significantly correlated with their domains and all domains are significantly correlated with the whole tool, which means that the tool have internal consistency validity.

Discriminate Validity was determined using Mann-Whitney Test. All (Z) values were significant at (0.01 – 0.05) levels.

To determine the reliability of the questionnaire, the researcher used Alpha Cronbach Formulas. The results showed that the questionnaire is highly reliable. In addition, Split Half Reliability Coefficients was also used to confirm the reliability. Results showed that all split half reliability coefficients are highly reliable.

To answer the second and the third questions of the study, the researcher used means, frequencies, standard deviations and percentages.

To answer the fourth question of the study, the researcher used (T Test) for two unequal independent samples,.

To answer the fifth question, the researcher used one way ANOVA (Analysis of Variance) and Scheffe Post-hoc test

# Chapter IV

# Findings

1. Introduction

2. Current trends for developing writing assessment

3. Existing assessment methods used in Gaza

4. Do existing methods match the current methods

5. Effects of teacher' gender on writing assessment

6. Effects of teacher' experience on writing assessment

7. Summary

# Chapter IV

# Findings

## 1. Introduction

This chapter introduces the findings and results of the study as they are statistically treated in addition to the discussions of the results with a reference to the five questions of the study. The researcher used different statistical forms such as means of frequencies and percentages to show the final collected data results. Tables and bar charts are also used to clarify and present these data.

## 2. Answer of Question 1

Question (1 ) **" What are the current trends for developing the assessment of writing in English language? "**

Based on more than 34 previous studies surveyed and theoretical framework, the researcher identified ten methods of assessment. Testing is considered as a traditional method of assessment and the rest of the nine methods are modern methods of assessment. The nine methods are:

1. Observation,
2. Self-assessment,
3. Peer-assessment,
4. Journals,
5. Interviews and conferences,
6. Portfolios,
7. Checklists,
8. Anecdotal records,
9. And rubrics

Although testing is regarded as a traditional method for assessing writing, it is needed to be employed together with other modern methods to achieve the utmost benefits of assessment. Every method of the ten was subdivided into different items. These items describe different procedures for applying each method. Here is the assessment methods which will discussed in full bellow.

**Testing:**

The first familiar method used for assessing student's writing is testing. It was subdivided into four items which includes different forms of tests. These items are:

a. Objective tests e.g. true/false, multiple choice, multiple-response and matching questions

b. Subjective tests e.g. extended-response questions and essays.

c. Diagnostic test

d. Quizzes

## 2.2 Observation

The second method used for assessing student's writing is observation. It is considered a modern method of assessment. Four items are included under this method. They summarize the procedures taken by teachers when observing their students during the writing tasks. These items are:

1. Observing students' attendance.

2. Observing behavior .e.g. their interaction.

3. Observing their performances during the writing tasks.

4. Writing short notes during a lesson, as students either work in groups or individually, or after a lesson.

### 2.3 Portfolios

Portfolios are the third method used for assessing student's writing. Four items are listed which includes the writing samples collected by teachers to be used for assessment purposes. These items are:

1. Students' best written work.

2. Students' reflection on why they chose the pieces they did.

3. Samples of previous writing tests and quizzes.

4. Self- assessment checklists filled by the students.


### 2.4 Self-assessment

Self- assessment is another method used for assessing students in different subjects which include also the assessment of students' writing. The following procedures are required to be carried out by teachers to apply the self-assessment method effectively.

1. Providing students with criteria check sheets (or have the class generate them) that specify exactly what constitutes a good product.

2. Training students gradually to assess their own work..

3. Training students to identify specific areas where they need more support and can seek help

4. Providing examples of other students' (anonymous) work to discuss in terms of its strengths and weaknesses.

5. Providing them with a description to the levels of quality of their work from excellent to poor.

### 2.5 Peer –assessment

Peer-assessment's method is similar to self-assessment and their procures are similar in some points. It permits learners to discover other students' styles, techniques, ideas and abilities. The following five items contain some procedures taken by teachers to apply the peer-assessment method effectively to assess students' writing.

1. Training students gradually to assess the work of other students Train students gradually to assess their own work..

2. Providing students with basic assessment criteria by which they assess their peers.

3. Using examples of work from anonymous learners as models of peer assessment.

4. Providing students with opportunities to assess their peers in situations in which their assessments do not affect their grades.

5. Encouraging student to share their works and discuss what might improve each other's work .

### 2.6 Journals

Journals is an effective method for assessing writing in which students have the freedom to write whatever they want and however they want. The next procurers include some items which the teachers guide their students to cover in the journal to be used for assessment purposes. These items are:

1. Documenting their expressions which they use when writing .

2. Documenting their feelings and reflections about the writing lessons and tasks.

3. Documenting their experiences and progress.

4. Documenting some of the difficulties they face in doing the task and how interesting or useful do they find the tasks they did

**2.7 Interviews and conferences**

In applying interviews and conferences for assessment, teachers are given the opportunity to interact personally with their students and learn about their different abilities, attitudes, strengths and weaknesses .To ensure the effectiveness of these methods for assessing writing, teachers need to consider the following procurers.

1. Identifying what students have learned and what they have yet to learn.

2. Assessing instructional changes or innovations

3. Giving individual feedback and clarify misconceptions.

4. Exploring individual differences in experiences and outcomes.

5. Recording information during or immediately following the communication.

**2.8 Rubrics**

Rubrics is an important method for assessment of writing by which students learn about the criteria on which their work will be judged. Teachers have to take into consideration the following procedures to gain the benefits of using this method .

1. Providing students with clear criteria against which they judge their own work.

2. Providing students with a description to the levels of quality from excellent to poor.

3. Sharing the descriptions of the scale with the students and ask for feedback so that each level is clearly understood by students.

4. Providing examples of students' work that illustrate each performance level.

5. Give students feedback clearly related to the assessment criteria.

**2.9 Anecdotal records**

The written observations that occur throughout the school day related to students' progress in learning are known as anecdotal records. The successful use of such a method entail the teachers to think about the following procedures when they employ it

1. Using anecdotal records to assess students participation and effort during the writing tasks.

2. Using anecdotal records to trace student's gradual progress toward attaining identified writing objectives.

3. Practicing using Anecdotal records systematically after each writing lesson or at the end of the day.

**2.10 Checklists**

The last method used for assessing student's writing is checklists. Two items are included to insure that checklists are used appropriately to assess students' writings.

1. Using checklists to guide students in self- and peer assessment.

2. Using checklists to record my observations during the writing tasks.

**3. Answer of question 2 : "What are the existing methods used by teachers for assessing English writing in 10<sup>th</sup> grade at Gaza schools from the teachers' perspective? ″**

The researcher used a questionnaire to find out the existing methods of assessing writing in English in 10<sup>th</sup> grade at Gaza governmental schools. The results of this questionnaire were analyzed statistically. Means, standard deviations and percentages were computed. The following tables analyze each of the ten methods statistically by

sorting out the items and showing the sum, mean and standard deviation. Then the same was done to the total questionnaire.

## 3.1. Testing

Table (12) presents the summation, means, standard deviation and the percentage of frequencies for each item in the testing domain. Out of the total percentage of frequencies, the item " Objective tests e.g. true/false , multiple choice, multiple-response and matching questions " got the highest score in a percentage of 90.1 %. The researcher thinks that the item got this high score because objective tests are easy to mark and scores are unaffected by the opinions of the scorers or examiners. They cover a number of educational objectives and students abilities rather than depending on rote learning. Additionally, the sampling that is covered is more representative of the content.

The" Subjective tests e.g. extended-response questions and essays"  item got the second score in a percentage of 79,8 while "Diagnostic test" was in the third rank in percentage of 4.96. The lowest item "Quizzes" is in a percentage of 35.6. The researcher thinks that is because they may create misunderstanding of assessment if seen as a threat and may increase teachers' workload. The total percentage of using test is 63.8% which is  rather high and indicates a large use of this method.

89

## Table ( 12 )

### Descriptive Statistics for Testing Domain

| N. | Items | Sum | Mean | S.D | Percentage | Sort |
|----|-------|-----|------|-----|------------|------|
| 1 | Objective tests e.g. true/false multiple choice, multiple-response and matching questions | 572 | 4.50 | 1.05 | 0.901 | 1 |
| 2 | Subjective tests e.g. extended-response questions and essays. | 507 | 3.99 | 1.05 | 0.798 | 2 |
| 3 | Diagnostic test | 315 | 2.48 | 1.33 | 0.496 | 3 |
| 4 | Quizzes | 226 | 1.78 | .99 | 0.356 | 4 |
| | Total Domai | 1620 | 12.76 | 2.02 | 0.638 | |

According to table (13),  9% of the teachers stated that they  sometimes do not resort to this method because they may not cover a wide range of content, while 22% attributed that to the subjectivity which may affect fair grading .

## Table(13)

### Reasons for not  using testing in assessment

| N. | Reason | Yes | | No | |
|----|--------|-----|-----|-----|-----|
| | | Number | % | Number | % |
| 1 | They may not sample a wide range of content | 11 | 9% | 116 | 91% |
| 2 | Subjectivity may affect fair grading | 28 | 22% | 99 | 78% |

## 3.2. Observation:

According to the results in table (14), the item "Observing behavior .e.g. their interaction" got the highest score in a percentage of 68.2 % . " Observing their performances during the writing tasks " item got the second store in a percentage of 67.2% while " Observing students' attendance" was in the third rank in percentage of 64.9 %. The lowest item " Writing short notes during a lesson, as students either work in groups or individually, or after a lesson" gets a percentage of 22.4%. The total percentage of using observation is 63.8% which is rather high and indicates a fair use of this method .

## Table (14 )

## Descriptive Statistics for Observation Domain

| N. | Items | Sum | Mean | S.D | Percentage | Sort |
|---|---|---|---|---|---|---|
| 2 | Observing behavior .e.g. their interaction | 433 | 3.41 | 1.12 | 0.682 | 1 |
| 3 | Observing their performances during the writing tasks | 427 | 3.36 | 1.27 | 0.672 | 2 |
| 1 | Observing students' attendance. | 412 | 3.24 | .92 | 0.649 | 3 |
| 4 | Writing short notes during a lesson, as students either work in groups or individually, or after a lesson. | 322 | 2.54 | 1.08 | 0.507 | 4 |
| | Total Domain | 1594 | 12.55 | 2.92 | 0.628 | |

As said by 27 % of the teachers, the low percentage of using this method is because they are time consuming, while 56% consider these methods not suiting the classes in Gaza which contain large number of students. Table (15) explains the results .

**Table (15)**

**Reasons for not using observation in assessment**

| N. | Reason | Yes | | No | |
|---|---|---|---|---|---|
| | | Number | % | Number | % |
| 1 | They are time consuming | 27 | 21% | 100 | 79% |
| 2 | They do not suit our classes which include large number of students | 71 | 56% | 56 | 44% |

## 3.3. Portfolio

As shown in table (16), the item " Students reflection on why they chose the pieces they did " got the highest score in a percentage of 31.8 %, " Students' best written work " item got the second store in a percentage of 28 while " Self- assessment checklists " was in the third rank in percentage of 27.1%. The lowest item " Samples of previous writing tests and quizzes " is in a percentage of 22.4%. The total percentage of using portfolio is 27.3% which is rather low and indicates modest use of this method.

## Table (16 )

## Descriptive Statistics for Portfolio Domain

| N. | Items | Sum | Mean | S.D | Percentage | Sort |
|---|---|---|---|---|---|---|
| 2 | Students reflection on why they chose the pieces they did | 202 | 1.59 | .85 | 0.318 | 1 |
| 1 | Students' best written work | 178 | 1.40 | .76 | 0.280 | 2 |
| 4 | Self- assessment checklists filled by the students | 172 | 1.35 | .69 | 0.271 | 3 |
| 3 | Samples of previous writing tests and quizzes | 142 | 1.12 | .32 | 0.224 | 4 |
| | Total Domain | 694 | 5.46 | 2.32 | 0.273 | |

Table (17) explains that 24% of the teachers attribute the low percentage of using this method to the increase in the workload when applying it , while 21% considered  it as time consuming.

## Table (17)

## Reasons for not  using observation in assessment

| N. | Reason | Yes | | No | |
|---|---|---|---|---|---|
| | | Number | % | Number | % |
| 1 | They  may increase my workload. | 96 | 76% | 31 | 24% |
| 2 | They are time- consuming | 100 | 79% | 27 | 21% |

## 3.4. Self assessment:

Table (18) shows that the item " Provide students with criteria check sheets (or have the class generate them) that specify exactly what constitutes a good product. " , in addition to the item " Train students gradually to assess their own work." got the highest score in a percentage of 23% .

The item " Provide them with a description to the levels of quality of their work from excellent to poor." was in the third rank in percentage of 22. 4%. In the forth rank was "Provide examples of other students' (anonymous) work to discuss in terms of its strengths and weaknesses" in a percentage of22 % . The lowest item "Train students  to identify specific areas where they." was in a percentage of 21.4%. The total percentage of using self assessment methods is 22.4% which is  rather low and indicates modest use of them by teachers.

# Table (18 )

## Descriptive Statistics for Self assessment Domain

| N. | Items | Sum | Mean | S.D | Percentage | Sort |
|---|---|---|---|---|---|---|
| 1 | Provide students with criteria check sheets (or have the class generate them) that specify exactly what constitutes a good product. | 146 | 1.15 | .36 | 0.230 | 1 |
| 2 | Train students gradually to assess their own work. | 146 | 1.15 | .38 | 0.230 | 2 |
| 5 | Provide them with a description to the levels of quality of their work from excellent to poor. | 142 | 1.12 | .39 | 0.224 | 3 |
| 4 | Provide examples of other students' (anonymous) work to discuss in terms of its strengths and weaknesses. | 140 | 1.10 | .35 | 0.220 | 4 |
| 3 | Train students to identify specific areas where they need more support and can seek help | 136 | 1.07 | .31 | 0.214 | 5 |
| | Total Domain | 710 | 5.59 | .89 | 0.224 | |

According to 62% of the teachers, the low percentage of using this method is due to the unfamiliarity of them , while 76% considered them to no suit cases with large number of students (see table 19).

## Table (19)

## Reasons for not using self- assessment method in assessment

| N. | Reason | Yes | | No | |
|---|---|---|---|---|---|
| | | Number | % | Number | % |
| 1 | I am not familiar with them | 79 | 62% | 48 | 38% |
| 2 | They do not suit our classes which have a large number of students | 97 | 76% | 30 | 24% |

## 3.5. Peer assessment

As revealed by table (20) , the three items " Provide students with basic assessment criteria by which they assess their peers " and" Provide students with opportunities to assess their peers in situations in which their assessments do not affect their grades " , in addition to the item " Encourage student to share their works and discuss what might improve each other's work " got the highest score in a percentage of 24.7%,

"Use examples of work from anonymous learners as models of peer assessment" was in the fourth rank in percentage of 22.7 % . The lowest item " Train students gradually to assess the work of other students " was in a percentage of 23. 8%. The total percentage of using these method is 23.8% which is rather low and indicates modest use oft hem by teachers .

## Table (20 )

## Descriptive Statistics for Peer assessment Domain

| N. | Items | Sum | Mean | S.D | Percentage | Sort |
|---|---|---|---|---|---|---|
| 2 | Provide students with basic assessment criteria by which they assess their peers | 157 | 1.24 | .65 | 0.247 | 1 |
| 4 | Provide students with opportunities to assess their peers in situations in which their assessments do not affect their grades | 157 | 1.24 | .65 | 0.247 | 1 |
| 5 | Encourage student to share their works and discuss what might improve each other's work | 157 | 1.24 | .65 | 0.247 | 1 |
| 3 | Use examples of work from anonymous learners as models of peer assessment | 144 | 1.13 | .46 | 0.227 | 4 |
| 1 | Train students gradually to assess the work of other students | 142 | 1.12 | .32 | 0.224 | 5 |
| | Total Domain | 757 | 5.96 | 2.29 | 0.238 | |

The results in table (21) show that 43%of the teachers stated that the low percentage of using this method is due to the Lack of support by the supervisors or others to apply these methods .A percentage of 29% of the teachers acknowledged that they have not received any type of training to apply them.

## Table (21 )

## Reasons for not using Peer assessment in assessment

| N. | Reason | Yes | | No | |
|---|---|---|---|---|---|
| | | Number | % | Number | % |
| 1 | Lack of support by the supervisors or others to apply these methods | 72 | 57% | 55 | 43% |
| 2 | I have not received any type of training to apply them | 90 | 71% | 37 | 29% |

## 3.6. Journals

As it clear in  table (22), the item " Document of their feelings and reflections about the writing lessons and tasks " got the highest score in a percentage of 24.7 %," Document of their experiences and progress "  item got the second store in a percentage of 23%..  The two items " Document of their expressions which they use when writing "  in addition to the item" Some of  the difficulties they face in doing the task and how interesting or useful do they find the tasks they did " were in the third rank   in a percentage of 22.4% .  The total percentage of using these method is 23.1% which is  rather low and indicates modest use of them by teachers .

# Table ( 22 )

## Descriptive Statistics for Journals Domain

| N. | Items | Sum | Mean | S.D | Percentage | Sort |
|---|---|---|---|---|---|---|
| 2 | Document of their feelings and reflections about the writing lessons and tasks | 157 | 1.24 | .65 | 0.247 | 1 |
| 3 | Document of their experiences and progress | 146 | 1.15 | .47 | 0.230 | 2 |
| 1 | Document of their expressions which they use when writing | 142 | 1.12 | .32 | 0.224 | 3 |
| 4 | Some of the difficulties they face in doing the task and how interesting or useful do they find the tasks they did | 142 | 1.12 | .32 | 0.224 | 3 |
| | Total Domain | 587 | 4.62 | 1.35 | 0.231 | |

Table (23) shows that 47 % of the teachers justified the low percentage of using this method by claiming that It can be difficult for them to determine assessment criteria, while 24% viewed them as a an increase of the workload which they can not afford.

# Table ( 23 )

**Reasons for not using Journals in assessment**

| N. | Reason | Yes | | No | |
|---|---|---|---|---|---|
| | | Number | % | Number | % |
| 1 | It can be difficult to determine assessment criteria | 67 | 53% | 60 | 47% |
| 2 | They may increase my workload | 97 | 76% | 30 | 24% |

## 3.7. Interviews & Conferences:

The descriptive Statistics for Interviews and Conferences in table (24) explains that the item " Give  individual feedback and clarify misconceptions   " got the highest score in a percentage of 31.8 % . " Assess instructional changes or innovations " item got the second score in a percentage of 28%  while " Identify what students have learned and what they have yet to learn. " was in the third rank in percentage of 27.1 %. In the forth rank was the item "Explore individual differences in experiences and outcomes " .  The lowest item   " Record  information during or immediately following the communication. " Was  in a percentage of 22.4 % . The total percentage of using these method is 23.7% which is  rather low *and* indicate modest use of them by teachers.

**Table (24 )**

**Descriptive Statistics for Interviews & Conferences Domain**

| N. | Items | Sum | Mean | S.D | Percentage | Sort |
|---|---|---|---|---|---|---|
| 3 | Give individual feedback and clarify misconceptions | 157 | 1.24 | .65 | 0.247 | 1 |
| 2 | Assess instructional changes or innovations | 151 | 1.19 | .56 | 0.238 | 2 |
| 1 | Identify what students have learned and what they have yet to learn. | 150 | 1.18 | .53 | 0.236 | 3 |
| 4 | Explore individual differences in experiences and outcomes | 148 | 1.17 | .56 | 0.233 | 4 |
| 5 | Record information during or immediately following the communication. | 147 | 1.16 | .56 | 0.231 | 5 |
| | Total Domain | 753 | 5.93 | 1.32 | 0.237 | |

Table (25) shows that 61% of the teachers attribute the low percentage of using this method is due to the lack of facilities to perform these methods , whereas 55% claimed that the have not received any training to try them for assessment purposes

**Table ( 25 )**

**Reasons for not using Interviews & Conferences in assessment**

| N. | Reason | Yes | | No | |
|---|---|---|---|---|---|
| | | Number | % | Number | % |
| 1 | There are no adequate facilities to perform these methods | 77 | 61% | 50 | 39% |
| 2 | I have not received any training to try them for assessment purposes | 70 | 55% | 57 | 45% |

## 3.8. Rubrics

The results in table (26) demonstrate that the item " Share the descriptions of the scale with the students and ask for feedback so that each level is clearly understood by students " got the highest score in a percentage of 31.8 %,

" Provide students with a description to the levels of quality from excellent to poor " item got the second store in a percentage of 28% . The three items " Provide students with clear criteria against which to judge their own work" , " Provide examples of students work that illustrate each performance level " and " Give students feedback clearly related to the assessment criteria." Were in the third rank in percentage of 27.1%. The total percentage of using these method is 23.8% which is rather low and indicate modest use of them by teachers

**Table (26 )**

## Descriptive Statistics for Rubrics Domain

| N. | Items | Sum | Mean | S.D | Percentage | Sort |
|---|---|---|---|---|---|---|
| 3 | Share the descriptions of the scale with the students and ask for feedback so that each level is clearly understood by students | 172 | 1.35 | .48 | 0.271 | 1 |
| 2 | Provide students with a description to the levels of quality from excellent to poor. | 157 | 1.24 | .43 | 0.247 | 2 |
| 1 | Provide students with clear criteria against which to judge their own work. | 142 | 1.12 | .32 | 0.224 | 3 |
| 4 | Provide examples of students work that illustrate each performance level. | 142 | 1.12 | .32 | 0.224 | 3 |
| 5 | Give students feedback clearly related to the assessment criteria. | 142 | 1.12 | .32 | 0.224 | 3 |
| | Total Domain | 755 | 5.94 | 1.63 | 0.238 | |

As said by 65 % of the teachers that the low percentage of using this method is due to the lack of training they received while 73% attributed that to the lack of facilities and teaching materials (see table 27).

## Table (27 )

**Reasons for not using rubrics in assessment**

| N. | Reason | Yes | | No | |
|---|---|---|---|---|---|
| | | Number | % | Number | % |
| 1 | I have not trained to use them | 82 | 65% | 45 | 35% |
| 2 | Lack of facilities or teaching materials | 93 | 73% | 34 | 27% |

## 3.9. Anecdotal records

As made obvious by table (28), the item " using anecdotal records to trace student's gradual progress toward attaining identified writing objectives. " got the highest score in a percentage of 31.8 %,

" practicing using Anecdotal records systematically after each writing lesson or at the end of the day. " item got the second store in a percentage of 29.4% . The lowest item " using anecdotal records to assess students participation and effort during the writing tasks. ", in a percentage of 24.7 %. The total percentage of using these method is 28.7% which is rather low *and* indicate modest use of them by teachers.

**Table ( 28 )**

### Descriptive Statistics for Anecdotal records Domain

| N. | Items | Sum | Mean | S.D | Percentage | Sort |
|---|---|---|---|---|---|---|
| 2 | Using anecdotal records to trace student's gradual progress toward attaining identified writing objectives. | 202 | 1.59 | .85 | 0.318 | 1 |
| 3 | Practicing using Anecdotal records systematically after each writing lesson or at the end of the day. | 187 | 1.47 | 1.30 | 0.294 | 2 |
| 1 | Using anecdotal records to assess students participation and effort during the writing tasks. | 157 | 1.24 | .43 | 0.247 | 3 |
| | Total Domain | 546 | 4.30 | 2.03 | 0.287 | |

As shown in table (29), time consuming received a percentage of 65% of the teachers' rejection to use methods, while 71% of the teachers considered them difficult to be applied in classes that contain large numbers of students.

**Table ( 29)**

**Reasons for not using Anecdotal records in assessment**

| N. | Reason | Yes | | No | |
|---|---|---|---|---|---|
| | | Number | % | Number | % |
| 1 | They are time consuming | 83 | 65% | 44 | 35% |
| 2 | I can not apply them on classes contain large number of students | 90 | 71% | 37 | 29% |

## 3.10. Checklists

From table (30) it is obvious that the item " using checklists to record my observations during the writing tasks " got the highest score in a percentage of 22.7 %  while the item "using checklists to guide students in self- and peer assessment. "   got the second score in a percentage of 22.4%. The total percentage of using these method is 22.5% which is  rather low and indicate modest use of them by teachers .

## Table (30  )

## Descriptive Statistics for Checklists Domain

| N. | Items | Sum | Mean | S.D | Percentage | Sort |
|---|---|---|---|---|---|---|
| 2 | Using checklists to record my observations during the writing tasks | 144 | 1.13 | .41 | 0.227 | 1 |
| 1 | Using checklists to guide students in self- and peer assessment. | 142 | 1.12 | .37 | 0.224 | 2 |
| | Total Domain | 286 | 2.25 | .52 | 0.225 | |

According to 76% of the teachers , the low percentage of using this method is due to the lack of training they received  while 80%  claimed that  they increase their workload ( see table 31)

**Table (31)**

**Reasons for not using Checklists in assessment**

| N. | Reason | Yes | | No | |
|---|---|---|---|---|---|
| | | Number | % | Number | % |
| 1 | I have not trained to use them | 97 | 76% | 30 | 24% |
| 2 | They  may increase my workload. | 101 | 80% | 26 | 20% |

## 3.11. Total Questionnaire:

The conclusion that can be drawn from table 32 that testing gained the highest score in a percentage of 63.8%. The next score was observation which got a percentage of 62.8%. The researcher found that testing and observation were the most widely used methods to assess the tenth grade students' writing in Gaza schools. The results confirm that most of teachers still rely on the traditional methods in assessing their students' writing. Race (2002)  stated that traditional written unseen exams still make up the lion share of assessment in education because they ensure the equality of the opportunity and plagiarism can be avoided. Brown (2004: 266)  pointed out that all teachers, whether they are aware of it or not ,observe their students in the classroom almost constantly.

Anecdotal records is in the third position with a percentage of 28.7%. The fourth is  portfolio which  gets 27.3 %. Rubrics  and  peer –assessment come in the fifth

position with a percentage of 23.8% . Interviews and conferences get a percentage of 23.7 % to be in the seventh position.  Journals get 23.1 % and checklists get the ninth position in a percentage of 22.5% and self- assessment gets the tenth position in a percentage of 22.4%.

**Table (32  )**

**Descriptive Statistics for Whole Questionnaire**

| N. | Domains | Sum | Mean | S.D | Percentae | Sort |
|----|---------|-----|------|-----|-----------|------|
| 1 | Testing | 1620 | 12.76 | 2.02 | 0.638 | 1 |
| 2 | Observation | 1594 | 12.55 | 2.92 | 0.628 | 2 |
| 3 | Portfolio | 694 | 5.46 | 2.32 | 0.273 | 4 |
| 4 | Self assessment | 710 | 5.59 | .89 | 0.224 | 10 |
| 5 | Peer assessment | 757 | 5.96 | 2.29 | 0.238 | 5 |
| 6 | Journals | 587 | 4.62 | 1.35 | 0.231 | 8 |
| 7 | Interviews & Conferences | 753 | 5.93 | 1.32 | 0.237 | 7 |
| 8 | Rubrics | 755 | 5.94 | 1.63 | 0.238 | 5 |
| 9 | Anecdotal records | 546 | 4.30 | 2.03 | 0.287 | 3 |
| 10 | Checklists | 286 | 2.25 | .52 | 0.225 | 9 |
| Total Questionnaire | | 8302 | 65.37 | 6.80 | 0.319 | |

## 4. Answer of Question 3 :"To what extent do the existing methods for assessing writing in English in 10th grade at Gaza schools match those of modern trends?"

The researcher recognizes that the  percentage of using testing which is the only traditional method of assessment is the highest with a percentage of 63.8%. The second score was for observation which is an alternative method of assessment with a percentage of 62.8. The other eight methods of assessment were rather low. The results confirm that the teachers still depend mostly on the traditional methods when assessing the writing of their students. So it is apparent that the existing methods for assessing  writing in English in $10^{th}$  grade at Gaza schools  do not match those of modern trends.

## 5. Answer of Question 4: "Is there a significant statistical difference at ($\alpha \leq$ 0.05) level related to the assessment methods used by male and female teachers?"

The researcher formulated the null hypothesis of the study which was stated as follows: "There is no statistically  significant difference at ($\alpha \leq 0.05$) level related to the assessment methods used by teachers according to their gender". The researcher used (T Test) for two unequal independent samples to test this hypothesis.  The following table describes the results:

## Table (33 )

## ( T Test) for the differences between males & Females

| N. | Domains | Gender | Mean | S.D | T value | Sig. |
|---|---|---|---|---|---|---|
| 1 | Testing | Male | 13.0615 | 1.87814 | 1.762 | No Sig. |
| | | Female | 12.4355 | 13.0615 | | |
| 2 | Observation | Male | 11.7538 | 2.86700 | 3.268 | Sig. at 0.01 |
| | | Female | 13.3871 | 2.76011 | | |
| 3 | Portfolio | Male | 4.2769 | .69614 | 6.937 | Sig. at 0.01 |
| | | Female | 6.7097 | 2.73644 | | |
| 4 | Self assessment | Male | 5.6615 | .94003 | 0.916 | No Sig. |
| | | Female | 5.5161 | .84430 | | |
| 5 | Peer assessment | Male | 6.6462 | 2.90167 | 3.616 | Sig. at 0.01 |
| | | Female | 5.2419 | .98656 | | |
| 6 | Journals | Male | 4.1692 | .51748 | 4.105 | Sig. at 0.01 |
| | | Female | 5.0968 | 1.74346 | | |
| 7 | Interviews & Conferences | Male | 6.0462 | 1.48356 | 1.021 | No Sig. |
| | | Female | 5.8065 | 1.12845 | | |
| 8 | Rubrics | Male | 5.8769 | 1.70942 | 0.478 | No Sig. |
| | | Female | 6.0161 | 1.56280 | | |
| 9 | Anecdotal records | Male | 3.3692 | .78201 | 5.958 | Sig. at 0.01 |
| | | Female | 5.2742 | 2.45062 | | |
| 10 | Checklists | Male | 2.2615 | .53843 | 0.212 | No Sig. |
| | | Female | 2.2419 | .50198 | | |
| Total Questionnaire | | Male | 64.4615 | 63.1231 | 4.040 | Sig. at 0.01 |
| | | Female | 70.5323 | 67.7258 | | |

According to table(33), it is obvious that there are no statistically significant differences at (α ≤ 0.05) level among teachers of English in Gaza schools  related to the assessment methods used by male and female teachers in the following : testing, self- assessment, interviews and conferences, rubrics and checklists. On the other hand, there is statistically significant differences at (α ≤ 0.01) level  with regards the following methods : observation, portfolio, peer-assessment, journals and anecdotal records. The results were in favor of females except  "peer –assessment" method which was in favor of males.

6. **Answer of Question 5: "Is there significant statistical differences at (α ≤ 0.05) level related to the assessment methods used by teachers according to their experience?"**

The researcher formulated the null hypothesis of the study which was stated as follows: "There is no statistically  significant difference at (α ≤ 0.05) level related to the assessment methods used by teachers according to their experience". The researcher used ANOVA (Analysis Of Variance) to test this hypothesis; the following tables describe the results:

## Table (34 ) Descriptive statistics for experience level

| Domain | Levels of Experience | N | Mean | Std. Deviation |
|---|---|---|---|---|
| Testing | less than 5 | 31 | 12.7419 | 2.25045 |
| | more 5 - 10 | 64 | 12.9844 | 2.10436 |
| | more than 10 | 32 | 12.3125 | 1.53323 |
| | Total | 127 | 12.7559 | 2.01863 |
| Portfolio | less than 5 | 31 | 6.4516 | 2.79092 |
| | more 5 - 10 | 64 | 5.1406 | 2.15191 |
| | more than 10 | 32 | 5.1563 | 1.88559 |
| | Total | 127 | 5.4646 | 2.31556 |
| Self assessment | less than 5 | 31 | 5.5484 | .80989 |
| | more 5 - 10 | 64 | 5.5313 | .92528 |
| | more than 10 | 32 | 5.7500 | .91581 |
| | Total | 127 | 5.5906 | .89402 |
| Peer assessment | less than 5 | 31 | 5.6774 | 1.77740 |
| | more 5 - 10 | 64 | 5.8750 | 2.26428 |
| | more than 10 | 32 | 6.4063 | 2.74578 |
| | Total | 127 | 5.9606 | 2.29008 |
| Journals | less than 5 | 31 | 5.1613 | 1.75303 |
| | more 5 - 10 | 64 | 4.4219 | 1.16571 |
| | more than 10 | 32 | 4.5000 | 1.13592 |
| | Total | 127 | 4.6220 | 1.35066 |
| Interviews & Conferences | less than 5 | 31 | 5.8710 | .95715 |
| | more 5 - 10 | 64 | 5.8125 | 1.44612 |
| | more than 10 | 32 | 6.2188 | 1.36155 |
| | Total | 127 | 5.9291 | 1.32246 |
| Rubrics | less than 5 | 31 | 5.4516 | .85005 |
| | more 5 - 10 | 64 | 5.9531 | 1.56783 |
| | more than 10 | 32 | 6.4063 | 2.16809 |
| | Total | 127 | 5.9449 | 1.63448 |
| Observation | less than 5 | 31 | 12.1613 | 3.19475 |
| | more 5 - 10 | 64 | 12.7969 | 2.76130 |
| | more than 10 | 32 | 12.4375 | 3.00470 |
| | Total | 127 | 12.5512 | 2.92148 |
| Anecdotal records | less than 5 | 31 | 5.1290 | 2.47308 |
| | more 5 - 10 | 64 | 4.2344 | 1.86652 |
| | more than 10 | 32 | 3.6250 | 1.62143 |
| | Total | 127 | 4.2992 | 2.03273 |
| Checklists | less than 5 | 31 | 2.1290 | .34078 |
| | more 5 - 10 | 64 | 2.3438 | .56957 |
| | more than 10 | 32 | 2.1875 | .53506 |
| | Total | 127 | 2.2520 | .51898 |
| Whole tool | less than 5 | 31 | 66.3226 | 8.47108 |
| | more 5 - 10 | 64 | 65.0938 | 6.21753 |
| | more than 10 | 32 | 65.0000 | 6.20614 |
| | Total | 127 | 65.3701 | 6.79613 |

## Table(35)(ANOVA) for the Differences between Teachers According to their Experience

| Domain | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| Testing | Between Groups | 9.638 | 2 | 4.819 | 1.186 | No Sig. |
| | Within Groups | 503.795 | 124 | 4.063 | | |
| | Total | 513.433 | 126 | | | |
| Portfolio | Between Groups | 39.960 | 2 | 19.980 | 3.898 | Sig. at 0.05 |
| | Within Groups | 635.631 | 124 | 5.126 | | |
| | Total | 675.591 | 126 | | | |
| Self assessment | Between Groups | 1.094 | 2 | .547 | .681 | No Sig. |
| | Within Groups | 99.615 | 124 | .803 | | |
| | Total | 100.709 | 126 | | | |
| Peer assessment | Between Groups | 9.310 | 2 | 4.655 | .886 | No Sig. |
| | Within Groups | 651.493 | 124 | 5.254 | | |
| | Total | 660.803 | 126 | | | |
| Journals | Between Groups | 12.055 | 2 | 6.028 | 3.432 | Sig. at 0.05 |
| | Within Groups | 217.803 | 124 | 1.756 | | |
| | Total | 229.858 | 126 | | | |
| Interviews & Conferences | Between Groups | 3.660 | 2 | 1.830 | 1.047 | No Sig. |
| | Within Groups | 216.703 | 124 | 1.748 | | |
| | Total | 220.362 | 126 | | | |
| Rubrics | Between Groups | 14.359 | 2 | 7.179 | 2.763 | No Sig. |
| | Within Groups | 322.256 | 124 | 2.599 | | |
| | Total | 336.614 | 126 | | | |
| Observation | Between Groups | 8.989 | 2 | 4.495 | .523 | No Sig. |
| | Within Groups | 1066.428 | 124 | 8.600 | | |
| | Total | 1075.417 | 126 | | | |
| Anecdotal records | Between Groups | 36.162 | 2 | 18.081 | 4.628 | Sig. at 0.05 |
| | Within Groups | 484.468 | 124 | 3.907 | | |
| | Total | 520.630 | 126 | | | |
| Checklists | Between Groups | 1.141 | 2 | .570 | 2.156 | No Sig. |
| | Within Groups | 32.796 | 124 | .264 | | |
| | Total | 33.937 | 126 | | | |
| Whole tool | Between Groups | 37.395 | 2 | 18.697 | .401 | No Sig. |
| | Within Groups | 5782.212 | 124 | 46.631 | | |
| | Total | 5819.606 | 126 | | | |

**Table (36)**

**Scheffe Post Hoc Test for multiple comparisons**

| Domains | (I) experience | less than 5 | 5 - 10 | more than 10 |
|---|---|---|---|---|
| Portfolio | less than 5 | | * | - |
| Portfolio | 5 - 10 | * | | - |
| Portfolio | more than 10 | - | - | |
| Journals | less than 5 | | * | - |
| Journals | 5 - 10 | * | | - |
| Journals | more than 10 | - | - | |
| Anecdotal records | less than 5 | | - | * |
| Anecdotal records | 5 - 10 | - | | - |
| Anecdotal records | more than 10 | * | - | |

* The mean difference is significant at the .05 level.

To check if there are significant differences between the used methods of assessment and the years of experience, ANOVA was used and results indicated a significant difference between teachers according to the years of experience in the area of portfolio, journals, and observation. However, no significant differences were found between teachers based on their years of experience in the other methods.

Scheffe Post-hoc analysis was conducted to examine where the differences exist. The post-hoc analysis revealed a significant difference at ($\alpha \leq 0.05$) between the teachers in the field of portfolio between the teachers of less than five years of experience and the teachers of more than five and less than ten years of experience. The results was in favor of the first group .

In the field of journal, a significant difference at ($\alpha \leq 0.05$) was found between the teachers of" less than five years of experience" and the teachers of" more than five and less than ten years of experience". The results was in favor of the first group .

In anecdotal records a significant difference at ($\alpha \leq 0.05$) was found between the teachers of" less than five years of experience" and the teachers of" more than ten years of experience". The results was in favor of the first group .

It is apparent that the result in the three domains were in favor of the teachers of " less than five years of experience "and the researcher think that this may because those teachers graduated recently and have some knowledge about this modern methods of assessment .

## 7. Summary

Chapter four presented the implementation of the questionnaire to evaluate the $10^{th}$ grade students' writing with regards to the current trends. The procedures discussed in this chapter were: how the questionnaire was conducted and applied to reveal the methods used by teachers to assess the students' writing through answering the four questions of the study.

The findings of the study drawn from the results show that there is a variance in using the ten methods of assessment. Even in each domain, there is a variance in presenting the items. There is a dominance of the testing and observation methods which respectively got 63.8% and 62.8%  compared with the other methods of assessment in which the highest was anecdotal records in a percentage of 28.7% and the lowest was self-assessment in a percentage of 22.4%.

The results that emerged from the questionnaire confirmed the importance of the study regarding the evaluation of the methods used in assessing the $10^{th}$ grade students' writing.

The final chapter will focus on consolidating the major findings of the study, drawing conclusions and making suggestions and recommendations based on these findings.

# Chapter V

## Discussion and recommendations

1. Introduction

2. Discussion

3. Suggestions

4. Recommendation for Further Studies

# Chapter V

# Discussion and Suggestions

## 1. Introduction

In the light of the educational literature reviewed and discussed throughout the study and the study findings, this chapter aims at discussing the results of the study, giving interpretation, providing valuable suggestions, and offering recommendation for further research.

## 2. Discussion

The study results reveal the lack of use of the modern methods of assessment and show the heavy reliance on the traditional method of assessment, testing. The only alternative method that got high scores was observation. The rest of other methods which the researcher classifies as modern ones got scores ranged from 28.7% for anecdotal records (the highest) to 22.4% for self- assessment (the lowest). These results reveal the rare use of these alternative methods.

It can be concluded that teachers in Gaza depend heavily on testing because they assess in order to rank students and not to identify individual capabilities and weaknesses. They assess not what they really want to assess, but what happens to be easy for them to assess. They do not give students a chance to improve themselves because they are administered after the learning process has taken place. Weeden et al (2002: 42) stated that the research evidence suggests that the assessment will have a negative impact on students learning if giving marks is overemphasized and giving useful advice and learning function are underemphasized. The further negative impact

happens when students are compared with one another because this will focus students' attention on competition rather than on personal improvement .

In this context, Yates (2005) said that teachers who perceive classroom assessment as testing may fail to understand the learning potentials and difficulties experienced by their students during the learning process. Information from assessment should help the teacher to discover areas where students have difficulties and can, therefore, be used to modify teaching methods and strategies in order to support students' learning.

The second traditional method included in the questionnaire was observation which got a percentage of 62.8%. Race (2003) assumed that observation is probably the most frequently used method in the classroom. Some negative points can be pointed out about observation. First it cannot probe the latent abilities of students especially in classes of large number of students. Students may be affected by some distracters which hinder them showing their real abilities. In this respect, Weeden et.al (2002: 136) affirmed that teachers need to remember the "invisible" students in their classes with whom they may have few direct interactions and about whom they feel they have less personal knowledge. Teacher's initial impression may affect the credibility of this method of assessment and teachers need to be conscious in developing ways of checking to make judgment accurate.

Portfolio assessment got a low percentage of 27.3% . which could be due to the fact that most teachers are unfamiliar with what is meant by portfolio, although some of them use some techniques in their classes that can be related to portfolio as a method of assessment. One problem voiced by teachers in applying this method was the time factor . Teachers were especially concerned with the amount of time this approach involves. A further concern was the lack of knowledge or training necessary

for implementing portfolio-based student assessment. In order for this form of assessment to be effective, teachers need to be trained in the various aspects of the approach. In this connection, Thomas et al (2005) stated that if teachers are simply given directions to implement portfolio assessment and mandated to do so without the proper preparation, it is doomed to failure. Just as teachers were trained in their specific discipline and/or content area, they should also be trained how to use the portfolio approach in assessment. Once trained and the plan implemented, teachers should also have follow-up training sessions through staff development. All participants should be kept abreast of developing research.

Self-assessment got a low percentage of 21.4% ,while peer-assessment got percentage of 23.8%. Confirming the pedagogical of these tools, Noonan & Duncan's (2005) state that teachers who use self- assessment and peer-assessment procurers operationalized an important component of "assessment for learning concept", that is, using assessment information to improve learning. They add that teachers who see the value of the reflective, self-analytic implement self-assessment. Similarly, teachers who understand the importance of collaborative, shared learning experiences, utilize peer assessment strategies. It is apparent from the questionnaire's result that some teachers do not include these two methods of assessment in their practice .They attributed that to their unawareness of these methods, lack of training and the large number of student  in each class .

Rubrics got a low percentage of 23.8% which indicates the limited use of them as  a method of assessment. Hedidi & du's study (2005) revealed the positive attitude of students towards using rubrics as a method of assessment. They  stated that rubrics helped them focus their efforts, produce work of higher quality, earn a better grade, and feel less anxious about an assignment. Andrade's study (2001) confirmed  the

119

benefit of rubrics in orienting students toward the criteria for writing and broadening students' conceptions of effective writing. Dorji (2000) offered some recommendation in his study included that : a) teachers need to be further trained to develop and use rubrics, b) teachers need to be trained to teach writing through the writing process approach, and c) schools need to adopt a new outlook towards writing and assessment. The previous recommendations are similar to the results revealed by the questionnaire in which teachers attributed their modest use of this method to the lack of training and the lack of facilities and teaching materials .

Teachers' responses to the questionnaire gave using Journal as a tool o assessment , a percentage of 23.1% . Snford's study (1997) approved that journal writing enabled students to explore their own ideas and feelings without being judged or censured. They were given opportunities to ask for assessments rather than having evaluation continually imposed upon them. Srimavin and Darasawang's study (2004) showed  that journal writing did not enable students to develop self-assessment. According to the researchers, that might be because there was no reaction from the tutor to what was written in the journals and the subjects treated the journal as an exercise where they answered the guiding questions rather than a place to assess their own learning. Teacher who do not use this method said that it may increase their workload and it can be difficult to them to determine the assessment criteria.

 Interviews and conferences provide teachers and students with numerous opportunities to assess progress.  At the heart of these types of assessment is the student and teacher working in collaboration in order to ascertain the progress the student has made. Nevertheless ,they got a low percentage of 23.7%. Lack of training and lack of facilities were the two main reasons as declared by the teachers for not using them for assessment .

Anecdotal records got a percentage of 28.7 %. Batstone (2004) stated that anecdotal records assessment provides teachers with an authentic tool to record observations in light of content standards. They facilitate assessment conversations between teachers, students and families. Lafi (2002) pointed out that checklists inform teachers about where their instruction has been successful and where students need assistance or further instruction. However, they got a percentage of 22.5%. Justification for the little use were concentrated on: lack of training, large number of students in the class, and the large amount of time and efforts the needed.

It is apparent from the previous studies the strong tendency towards the adopting of new methods of assessment. However, the teachers in Gaza schools still depend largely on the traditional methods of assessment . So, it is important to help new and experienced teachers to become more familiar with alternative assessments through classes, workshops and other professional development activities.

Some factors which hinder the use of the modern methods of assessment include: lack of teacher's knowledge, lack of skills and experience, lack of teacher's support, large classes, inadequate teaching and learning resources and lack of proper and sufficient training.

The researcher thinks that the most important factor to be considered is lack of knowledge, skills , experience and sufficient training . It was noticed during my visits to Gaza schools that most teachers do not know these contemporary methods of assessment. They may implement some procedures but they do not adopt them as a strategy and part of clear specific assessment plan that intends to promote writing skills. Others may implement them to increase students motivation or to change the class atmosphere.

It is obvious that there is no common strategy to adopt these methods of assessment at the level of the Ministry of Education, supervisors or principals. Hence, the  first step should be taken by the Ministry of Education to adopt these methods of assessment and to adapt the most suitable for our context. Then it should provide teachers with training courses to familiarize them with these methods and how to use them effectively.

A strong pedagogical and personal belief should be formed by teachers  in trying new ways of alternative assessment in order to create their willingness to implement alternatives methods of assessment  which suit Gaza classrooms. It  is important to recognize that small initial steps are necessary to make the change into assessment. This change needs support of a wide range of people. It does not happen overnight and teachers need to be able to experiment and share ideas and find out what work for them in their context.

Concerning the  assessment methods used by male and female teachers, results indicate that there is statistically significant differences at ($\alpha \leq 0.01$) level   with regards the following methods : observation, portfolio, peer-assessment, journals and anecdotal records .The results were in favor of females with the exception of  "peer – assessment" method  which was in favor of males.

The researcher thinks that this result is due to two reasons . Firstly, in Gaza schools female teachers teach female students and male teachers mostly teach male students. Female students respond more positively to their teachers' instructions and directions which enhance their teachers' motive towards the application of new methods of assessment. Zhongping, (2001) stated that female one-sidedly emphasize hardworking, obedience and disciplines in their study.

Secondly, in our society, men are mainly responsible for their family support and the low income may force them to get another job to meet the demands of their family. So, this further load may make the male teachers think less in changing and developing their practice.

The study revealed a significant difference between teachers according to the years of experience in the area of portfolio, journals, and observation. The results were in favorer of the less experienced teachers .

The researcher thinks that applying these methods of assessment is not related to the experience because they are new methods and the teacher who spent more years in teaching may be accustomed to the traditional methods of assessment and not aware of most of these methods. On the other hand, the teachers who are less experienced may be graduated recently and may come through these methods during their study in the university .

## 3. Suggestions

In the light of the results and the information obtained by the research, some important suggestions should be voiced to gain the fruits of the research.  A change in the education system should be taken to move towards integrating the alternative methods of assessment in the teaching process. They need to be incorporated systematically into teaching strategies and practices at all levels. This process of change should involve  teachers, principals, university faculties, parents, and Ministry of Education to work collaboratively to achieve this purpose.

The process of change starts by changing the beliefs. Accordingly, a new strong belief should by embraced by all members of the educational system to pave the way towards adopting modified method of alternative assessment that suit Gaza schools.

Two important concepts support the endeavor to put this change in action . The first is the learner-centered concept which places the student at the center of education and places the responsibility for learning on the student, while the instructor assumes responsibility for facilitating the student's education. It seeks to meet the individual needs of a broad range of learners who have different learning style, skills and backgrounds. Different learning styles stress the need for employing a variety of assessment methods. This means that the focus is on what Richards (2004) called the process of learning rather than the product, that the content and the teacher adapt to the students rather than expecting the students to adapt to the content. "Students will need to recognize when information is required, how to locate and retrieve information, and how to analyze the information to become useful"( Haugen 1998 ).

The second concept is  the "assessment for learning", as it was called by Richards (2004)  .It means that assessment becomes part of the learning process. The tools used

for assessment become tools for learning . They provide information to be used as feedback to modify the teaching and learning activities in which teachers and students are engaged.

The main role in change should be initiated by Palestinian Ministry of Education. It should adopt the new methods of assessment as a strategy and take the necessary steps towards the implementation of the most suitable of these methods .

A starting point should be the formation of a clear vision by the experts and policymakers. This vision leads to a rationale contains a brief description of the reasons of using alternative methods of assessment and the nature of them.

In order to develop the rationale , a plan should be formed to start implementing the most suitable of these methods. The plan should make use of available literature on the topic, published material on the topic, consultation with specialists in the area and consultation with teachers familiar with the topic.

Next step is to try this plan with representatives classes or schools before it is made available for wider use in order to find problems or faults in it that can be identified before it is used more widely.

Proper and sufficient training should be provided to student teachers. Courses should be designed to familiarize teachers with these methods and increase their effectiveness in using the most suitable in our context. Academic members of staff should mentor teachers to guarantee the implementation of these methods in the schools in which teachers are trained.

At the same time, professional support and guidance through specialized consultants should be made available to teachers working in the field until they gain confidence; at least during the initial period. Teachers must also have access to resource materials in the form of books and articles to read and opportunities to share

and discuss experiences through workshops, journals, and school magazines. Besides, teachers should be encouraged to conduct action research to check the suitability of a given methods of assessment to Gaza classes.

Once teachers are trained and the plan implemented, there should also be follow-up training sessions to give them the opportunity for further professional development. They need to find feedback to know if they are doing well and when there are problems with their performance. Acknowledge of good service should be given to teachers who perform well. Additionally, help lines should be opened to teachers in case they have difficulties in implementing these methods and using assessment materials. This may include opening a website for teaches to provide them with any consultations and support in how to apply a given method. It further can provide them with successful experience made by teachers in this field.

It is worth pointing out that teachers are the key factor in the process of change However, they can only be expected to make these changes if they are encouraged to do so within a supportive environment including all participants in the teaching process like principals, supervisors, faculties of education, Ministry of Education and parents.

Principals and supervisors have important role in providing teachers with the ongoing support. They should provide teachers with all facilities and teaching materials that can help them in applying the new type of assessment . The successful experience should be disseminated to other teachers and the teachers who succeed in their mission should be recompensed. Systematic visits of the principals and supervisors have the role of guiding and helping students in their practice with regards to this topic. To give more support for the implementation of the new methods, principals and supervisors should plan to invite university professors of English to

visit schools to hold discussion with the teachers of English concerning the recent and valid methods of assessment.

Faculties of Education have an important role  in facilitating and supporting teachers' work .They can train student teachers of English to  use these alternative methods of assessment and provide them with the most modern methods and the most suitable methods for Gaza schools. Teachers can be  invited to attend or participate in conferences to keep in touch with the recent  and appropriate methods of assessment. Moreover, they should supply the schools and Ministry of Education with empirical researches about the new trends in language assessment . A framework to exploit technology  should be offered to teachers  to make better use of these methods of assessment. For example, students can send their journals using the email to the class homepage where the teacher can give them feedback.

Parental or community support for such new and unfamiliar methods of assessment is essential. Parents  need to get accustomed to the change which could be difficult for them to accept or adjust to without considerable effort to educate them as to the nature and advantages of using such methods.

It is worth mentioning that there is a need to develop a culture of continuous or ongoing professional development. To insure  effective implementation of these new methods of assessment, cooperation should happen among all the participants in the assessment process ; teachers , school principals , supervisors and parents .

## Recommendation for further studies

The researcher offers the following recommendations for further studies:

1. Carrying out other studies to explore the effectiveness  of implementing one or some  of these alternative methods for assessing writing in Gaza schools.

2-. Doing other studies to explore the extent to which these alternative methods of assessment are used to assess students in other skills other than writing such as: listening, reading and speaking.

3. Carrying out further studies to investigate the use of alternative methods of assessment   in other institutions like UNRWA schools, universities and colleges.

# References

Abdel-Wahab, S (2002) *Portfolio Assessment : A qualitative Investigations of portfolio Self-Assessment Practice in An Intermediate EFL Classroom, Saudi Arabia.* [Online] Unpublished doctorial of education dissertation. The Ohio State University
Retrieved on the 3rd August , 2006 from
http://www.ohiolink.edu/etd/sendpdf.cgi/Alabdelwahab%2C%20Sharif%20Q.pdf?osu1037375641

Allan, c ( 1995) *Using Rubrics to Assess Learner through Service in Maini* [Online] Oxford Hills Comprehensive High School.
Retrieved on the 3rd August , 2006 from
http://www.vermontcommunityworks.org/cwpublications/slassessguide/AssessGdChpt4.pdf

Allen, D (1998) *Assessing Student learning from Grading to Understanding.* Teachers College Press: New York.

Allen, M (2004) The *Use of Scoring Rubrics for Assessment and Teaching* [Online] Director, California State University Institute for Teaching and Learning
Retrieved on the 3rd August , 2006 from
http://cai.cc.ca.us/Resources/RubricsByMaryAllen.doc

Alverno College Faculty. (2006). *Student assessment-as-learning at Alverno College.* Milwaukee, WI: Alverno College Institute.
Retrieved on the 23rd August , 2006 from
http://depts.alverno.edu/saal/saframe.html

American Evaluation Association (2007*) Evaluation Definition: What is Evaluation ?* [Online] American Evaluation Association.
Retrieved on the 3rd August , 2006 from
http://www.evaluationwiki.org/index.php/Evaluation_Definition

Andrade, H (1997) *Rubric and Self Ass*essment Project.[Online] The Edna McConnell Clark Foundation.
Retrieved on the 9th August , 2006 from
http://www.pz.harvard.edu/Research/RubricSelf.html.

Andrade, H (2001) *The effect of Instructional Rubrics on Learning to Write* [Online] current Issue in Education volume 4 number4
Retrieved on the 4th August , 2006 from
http://cie.asu.edu/volume4/number4/index.html

Apple, M & Shimo, E (2003) *Learners to Teacher : Portfolios, Please Perception of Portfolio Assessment in EFL Classrooms* [Online] Tokyo Keizai University
Retrieved on the 4th August , 2006 from
http://jalt.org/pansig/2004/HTML/AppleShimo.html

Aschbacher, p et al (1995) *Los Angeles Learning Center Alternative Assessment Guidebook* [Online] National Center for Research on Evaluation, standards, and Student Testing (CRESST)
Retrieved on the 12th August , 2006 from
http://www.cse.ucla.edu/products/guidebooks/gbthree.pdf

Assessment Companion (2001) *Assessment* [Online] Ontario Ministry of Education, April 2001
Retrieved on the 12th May , 2006 from
http://edu.queensu.ca/resources/resources/downloads/companions/assesg.pdf

Bacha (2001) *Testing Writing in The EFL Classroom : Student Expectations* [Online] Bureau of Educational and Cultural Affairs
Retrieved on the 12th May , 2006 from
http://exchange.state.gov/forum/vol/vol140/no2/p14.html

Batstone, P (2004) *Focused Anecdotal Records Assessment : A lot for Standards-based, authentic assessment* [Online]
Retrieved on the 12th August , 2006 from
http://www.learner.org/channel/workshops/teachreading35/pdf/anectodal_records.pdf

Borich, G (4004) *Effective Teaching Methods* Pearson Prentice Hall: New Jersey

Brooks, L L (1999) Adult ESL Student Attitudes towards Performance-Based Assessment [online] Department of Curriculum, Teaching and Learning Ontario Institute for Studies in Education for the University of Toronto.
Retrieved on 8th August,2006 from
http://www.collectionscanada.gc.ca/obj/s4/f2/dsk1/tape7/PQDD_0002/MQ45956.pdf

Brown, D (2004) Language Assessment Principle and Classroom Practice. Pearson Education, Ink

Brown, S et al (1994) Involving Students in the Assessment Process [Online]  The Oxford centre for Staff and Learning Development
Retrieved on the 4th August , 2006 from
http://www.city.londonmet.ac.uk/deliberations/ocsd-pubs/div-ass5.html

Chen , Y & Wei , Y (2003*) Supporting Chinese Learners of English to Implement Self – assessment in L2 Writing*  [Online] Proceedings of the Independent Learning Conference 2003
Retrieved on the 4th August , 2006 from
http://www.independentlearning.org/ila03/ila03_wei_and_chen.pdf

Chen, L(2002) *Taiwanese Junior High School English Teacher ' Perception of the Washback effects of the Basic Competence Test in English* [Online] The Ohio State University
Retrieved on the 4th August , 2006 from
http://www.ohiolink.edu/etd/view.cgi?acc_num=osu1039116364

Cobbine, J (1995) *Effective Use of Student Journal Writing[Online Eric Claringhouse on Reading, English , and Communication* Digest # 99
Retrieved on the 6th August , 2006 from
 http://reading.india.edu/ieo/digest/d99.html

Coombe, C (20002) *Self-Assessment in Language Testing :reliability and Validity Issue* [Online] Higher colleges of Technology, UAE
Retrieved  in the 12th November , 2006 from
http://www3.telus.net/linguisticsissues/selfassess.html

Connor, G et al (2001) *Does Writing Matter? Assessing the Impact of Daily Essay Quizzes in Enhancing Student learning*
Retrieved  on the 4th August , 2006 from
http://wac.colostate.edu/llad/v4n1/connor-greene_murdoch.pdf

Grehan, K & Curfman, M (1999) *The Effect on Performance of Rapid Feedback on State Writing Assessment*
Retrieved  in the 24th March , 2006 from
http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/15/f0/75.pdf

Dorji, T (2000) *Rubrics Assessment as an instructional Illumination of Bhutanese Student writing* [Online] unpublished MA thesis. University of New Brunswick.
Retrieved on 18th August, 2006from
http://dspace.hil.unb.ca:8080/bitstream/handle/1882/339/MQ62123.pdf?sequence=1

East. H  (2006) *Group, peer and self assessment* [Online] the UK centre for legal Education.
Retrieved on 22nd June 2006 from
www.ukcle.ac.uk/resources/assessment/group.html - 19k

Echevarria, J et al (2004) *Making Content Comprehensible For English Learners The SIOP Model* Pearson Education Inc: USA

Fall, et al (1997) *Group Discussion and large- scale language art assessment: Effects on students comprehension* [Online] CRESST/ University of California Los Angeles.
 Retrieved on 8th August, 2006from
http://aer.sagepub.com/cgi/content/abstract/37/4/911

Farr, R and Tone, B (1994) *Theory Meets Practice in Language arts Assessment* [Online] ERIC Educational Reports.
Retrieved on 22nd  August, 2006from
http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/15/7b/5c.pdf

Flucher, G and Davison, F (2007) *Language Testing and Assessment* , Routledge Taylor & Francis Group.

Foundation  Coalition (2007*) Assessment and Evaluation*
Retrieved on 22ⁿᵈ  August, 2006from
http://www.foundationcoalition.org/home/keycomponents/assessment_evaluation.html

Gaith, G (2002) *Using Cooperative Learning to Facilitate Alternative assessment English*. [Online] Teaching forum
Retrieved on 4ᵗʰ August, 2006from
http://exchanges.state.gov/forum/vols/vol40/no3/p26.htm

Gansle .et. al (2005) *Elementary School Teachers perceptions of Curriculum- Based Measures of written Expression* [Online] Practical Assessment, Research & Evaluation Journal Volume 11, Number 5. March 2006
Retrieved on 5ᵗʰ  July,2006 from
http://pareonline.net/getvn.asp?v=11&n=5

General Administration of Curricula (1999) *English Language Curriculum For Public Schools Grade_ 12  Ministry of Education Palestinian National Authority*

Genesee, F. and Upshur, J. (1996) *Classroom-based Evaluation in Second Language Education*. Cambridge University Press: Cambridge.

Gearhart et al (1993) *Writing Portfolios at Elementary Level: A study of Methods for Writing Assessment*. Eric  ED 344 900
Retrieved  on the 4th August , 2006 from
http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/24/02/86.pdf

Greece, C (2002) *EFL Writing: Product and Process* [Online Karen's Linguistics Issue
Retrieved  on the 5ᵗʰ  August , 2007 from
http://www3.telus.net/linguisticsissues

Hamer. J et. al (2005) A *Method of Automatic Grade Calibration in Peer Assessment* [Online] Department of Computer Science University of Auckalnd.
Retrieved on 18ᵗʰ August,2006 form
http://www.celt.iastate.edu/teaching/syllabi.html

Haugen, L (1998) *Learning-Centered Syllabi Workshop*[Online] Center for excellent in learning and teaching
Retrieved on 22nd  August, 2006from
http://www.celt.iastate.edu/teaching/syllabi.htm

Heidi, A. and Du, Y. (2005) *Student Perspectives on Rubric-Referenced Assessment.* [Online] Practical Assessment, Research & Evaluation (Volume 10 Number 3, April 2005 ISSN 1531- 7714) University at Albany.
Retrieved on 19th February,2006 from
http://Pareoline.net/getvn.asp?v=10&n=3

Hirsch. S (1996) A comparative study "Rubric" [Online] San Diego, California
Retrieved on 16th may, 2006 from
http://edweb.sdsu.edu/tritoin/SDBiarritz/rubic.html

Heinz P (2004) *Towards Enhanced Second Language Reading /Comprehension Assessment: Computerized Versus Manual Scoring of Written Recall Protocols* [Online] reading in Foreign Language. Volume 16.Nummber2,October 2004
Retrieved on 8th January, 2006 from
http://nflcr.hawaii.edu/rfl/october2004/heinz/heinz.html

Herman, J. et al (1997) *American Student Perspectives on Alternative Assessment: Do they Know it's Different?* [Online] National Center for Research on Evaluation, Standards, and Student testing (CRESST) CSE Technical Report 439. University of California, Los Angeles.
Retrieved on 8th August, 2006 form
http://www.cse.ucla.edu/Reports/TECH439.pdf

Herman and Gearhart (1996) *Issue in Portfolio Assessment: The scorability of narrative collections* [Online] National Center for Research on Evaluation, Standards and Student Testing (CRESST) University of California, Los Angles
Retrieved  on May 4th 2006 from
http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED462388

Isaacs. G (1996) *Brief Briefing: Pear and self assessment* [Online] the teaching and Educational Development Institute The University of Queensland.
Retrieved on 16th May, 2006 from
http://www.tedi.uq.edu.au/conferences/A confl/papers/isacs.html

Lafi, M (2002) *Alternative assessment* [Online] Republic of Tunisia Ministry of Education
Retrieved on 23rd  September, 2006 from
http://www.anglais.edunet.tn/tutorials/trainl/pdf/alternatives_assessment.pdf

Lake. C (1998) *Defining Assessment* [Online] CESN National Web
Retrieved on 18th August 2006 from
http://www.essentialschools.org

Kennesaw State University(2006) *Assessment Rubrics*
   Retrieved on 16th April 2006 from
   http://edtech.kennesaw.edu/intech/rubrics.htm


Kenworthy (20060 *An investigation of student self- assessment and writing Progress in An English competition course at Aattimura University Ambon, Andonisia* [Online] Unpublished MA thesis. Simon Fraser University.
   Retrieved on May 4th 2006 from
   http://amicus.collectionscandad.cs/s4bin/manin/BasicSearch?1=0&1_ef1=1&id=45073.155136&v=0&coll=18&1v1=4&dt=PD+|language|+AnND+PD+|testin|&rsn=pdfsearch.tmp&rpw=1&vo=1

Knight, P and Yorke, M (2003) *Assessment , Learning and Employability* Open
   University Press Glasgow: Glasgow


Marzano, R. J., Pickering, D. J., & McTighe, J. (1993). *Assessing student outcomes:*

   *Performance assessment using the Dimensions of Learning model.* Alexandria,
   VA: Association for Supervision and Curriculum Development.


Maxwell. G (1999) *Teacher Observation in Student Assessment* [Online] School of
   Education, The University of Queensland.
   Retrieved on 23rd September 2006 from
   http://www.qsa.qld.edu.au/yrs1to10/oia/assess/observ.pdf

Min. B (2002) *Relevant Assessment Strategies for online Colleges & Universities*
   [Online] USDLA journal Volume. 16: number. 2
   Retrieved on 18th August 2006 from.
   http://www.usdla.org/html/journal/FEBO2 Issue/aricle04.html

Moskal B. (2003) *Recommendations for Developing Classroom Performance Assessments and scoring Rubrics* [Online] Practical Assessment, Research & Evaluation
   Retrieved on 18th May 2006 from
   http://pareonline.net/getvn.asp?v=8&n=14


Online USDLA journal Volume. 16: number. 2
   Retrieved on 11th April, 2006 from
   http://www.usdla.org/html/journal/FEBO2_issue/article04.html

Nakamura Y (2002) *A comparison of Holistic and analytic scoring Methods in the assessment of writing* [Online] Tokyo Keizai University
   Retrieved on 12th April, 2006 from
   http://jalt.org/pansing/2004/HTML/Nakamura.htm

NCLRC (2004) *Assessing learning, Peer and Self Assessment* [Online] National Capital Language Resource Center 2006.
Retrieved on 18th August, 2006 from
http://www.nclrc.org/essentials/assessing/peereval.htm

Noonan, B. & Duncan, R. (2005) *Peer and Self-Assessment in High Schools*. [Online] Practical Assessment, Research & Evaluation (volume 10 number 17, November 2005, ISSN 1531-7714) University of Sakatchewan.
Retrieved on 21st August, 2006 from
http://pateonline.net/getvn.asp?v=10&n=17

O'Mally (1992) *Performance and Portfolio Assessment For Language Minority Students* NCBE Program Information Guide Series, Number 9, Spring 1992.
Retrieved on 18th May, 2006 from
http://www.ncela.gwu.edu/pubs/pigs/pig9.htm

Paulson. H (2002) *Assessment Tools "Anecdotal Records"* [Online] George Washington University.
Retrieved on 25th August, 2006 from
http://home.gwu.edu/~anniehp/portfolio/anecdotal_Records.htm

Park university, (2003) *Incorporating Portfolio Assessment* [Online] Park university
Retrieved on 30th August, 2006 from
http://www.park.edu/cetl/quicktips/porltolio.html

Patterson, A. & Bellaby, G. (2002) *An Initial Experiment with Computer Aided Assessment in level One System Analysis*. [Online] 3rd Annual LTSN-ICS Conference, Loughborough University    2002 LTSN Center for Information and Computer Sciences.
Retrieved on 18th August, 2006 from
http:///www.ics.heacademy,ac.uk/Events/conf/2002/apatterson.pdf

PGCPS (2006) *What Are Some Different Types of Portfolios?* [Online] Price George's School Country Public School.
Retrieved on 7th August, 2006 from
http://www.pgcps.org/~elc/portfolio2.htm

Plimmer, B. (2000) *A Case Study of Portfolio Assessment in a Computer Programming Course.* [Online] Proceedings of the NACCQ 2000, Manukau Institute of Technology Manukau City, New Zealand
Retrieved on 18th May, 2006 from
http://www.in-site.co.nz/misc_links/papers/plimmer279.pdf

Quessenberry, L. et al (1999) *Assessment of the writing Component within a University General Education Program.* [Online] Academic writing Navigator. Clarion University.
Retrieved on 18th August, 2006 from
http://wac.colostaqte.edu/aw/articles/quesenbery2000/quesenbery2000.pdf

Reid, L. (2003) *'Now I Know what to do!' Teachers talking about writing assessment.* [Online] Scottish Educational Research Association Annual Conference (perth, Scotland, 27-29 November 2003).
Retrieved in 8th February, 2006 from
http://www.leeds.ac.uk/educol/doucuments/00003486.htm

Richards, J (1996) *Classroom -Based Evaluation in Second Language Education,* Cambridge University Press: Cambridge.

Richards, J and Renandya, W (2002) *Methodology in Language Teaching.* Cambridge: Cambridge University Press.

Russell, M. & Haney, w. (1997) *Testing writing on Computers: An Experiment Comparing Student Performance on tests conducted via computer and via Paper-and-Pencil.* [Online] Education policy Analysis Archives (Volume 5 number 3: January 1997, ISSN 1086-2341).
Retrieved on 8th August, 2006 from
http://epaa.asu.edu/epaa/v5n3.html

Shaaban, K (2001) *Assessment of Young Learners* [Online] Bureau of Education and Cultural Affairs.
Retrieved on 18th March, 2006 from
http://exchanges.state.gov/forum/vols/vol39/no4/p16.htm

Snford, K (1997) *Caught in the Nixed Message of Evaluation: Teachers Evaluation Student writing Student Learning Assessment Implementation Team (2002) Report on Assessment of Writing through a Common Writing Assignment 2001-2002*[online] Capital: Community college.
Retrieved on 7th August, 2006 from
http://www.ccc.comment.edu/slat/cwa_summary.htm

Sook, K. (2003) *The Types of Speaking Assessment Tasks Used Korean Junior Secondary School English Teachers.* [Online] Asian EFL Journal: L2 Language assessment in the Korean Classroom (December, 2003)
Retrieved on 18th March, 2006 from
http://www.asian-efl-journal.com/dec_03_gl.pdf

Srimaving, W. and Darasawang, P. (2004) *Developing self- assessment through journal writing.* (online) Proceedings of the Independent Learning Conference 2003 (Published 20 September 2004)
Retrieved on 7th February, 2006 from
http://www.independentlearning.org/ila03/ila03_srimavin_and_pornapit.pdf

Tasagari, D (2004) *Is there life beyond language testing? An introduction to*

*alternative language assessment.* Center For Research In Language Education, Lancaster University
Retrieved on 7th February, 2006 from
http://www.ling.lancs.ac.uk/groups/crile/docs/crile58tsagari.pdf


The Natinal Center for Fail & Open Testing (2006) *How Standardized Testing Damages Education* [Online] The National Center for Fair & Open Testing
Retrieved on 18th February, 2006 from
http://www.fairtest.org/facts/howharm.htm

The Office of University Evaluation (2002) *General Studies Literacy Assessment.* [Online] The ASU General Studies Council and the Office of University Evolution.
Retrieved on 8th May, 2006 from
http://www.asu.edu/oue/old/reports/012_gs_literacy_report.pdf

Veronesi (1997) *A Case study of Alternative Assessment: Student, Teacher, and Observer Perceptions in A ninth Grade Biology Classroom* [Online] state University of New York at Brockport.
Retrieved on 6th May, 2006 from
http://www.ed.psu.edu/ci/journals/97pap39.htm



Vose .J (1997) *Assessment Strategies to Inform Science and Mathematics Instruction* [Online] Northwest Regional Educational Laboratory.
Retrieved on 21st August, 2006 from
http://www.nwrel.org/msec/images/resources/justgood/06.97.pdf

Wangchuk, K. (2000) *The Feasibility of Rubrics and Portfolio Assessment for Use in Continuous Assessment in Bhutan.* [Online] Unpublished Master of Education thesis. The University of New Brunswich.
Retrieved on 6th June, 2006 from
http://www.unbf.ca/education/bhutanproject/documents/LibraryLists.pdf

Yates, K. (2005) *Perceived Effectiveness of Assessments Used in Online Courses in Western North Carolina Community Colleges.* [Online] Unpublished doctorial of education dissertation. University of East Tennessee State University.
Retrieved on 21st August, 2006 from
http:///etd-submit.etsu.edu/etd/theses/available/etd-1102105-213208/unrestricted/YastesK110905f.pdf

Weimer (2002) *Learner-Centered Teaching Does Not Deny The Importance of Grades* [Online]
Retrieved on 8th August, 2006 from
http://tep.uoregon.edu/workshops/teachertraining/learnercentered/assessing/texdocs/lc_assessing.pdf

Worley. M (2001) *Alternative Assessment: Methods To Make Learning more Meaningful* [Online] College of Education, Armstrong Atlantic state University, Savannah
Retrieved on 22nd March, 2006 from
http://k12s.phast.umass.edu/stemtec/pathways/proceedings/papers/worley-p.doc

Yueming (2006) *Exploring the Establishment of Scoring Criteria for Writing Ability in A second Language: the Influence of Background Factors on Variability in the Decision-Making Processes of four Experienced Raters of ESL Compositions.* [Online] Ontario Institute for Studies in Education, University of Toronto.
Retrieved on 4th February, 2006 from
http://www.collectionscandada.ca/obj/s4/f2/dskl/tape3/PQDD_0019/MQ49745.pdf

Zhongping . H (2001) *How the Trend of Girls Achieving Better Result Than Boys Emerges in China's Schooling* [Online] AARE 2001 International Education Research Conference Fremantle. Australia.
Retrieved on 23rd June, 2006 from
http://www.aare.edu.au/01pap/hu01329.htm

# Appendix I

## A List of Experts

| | | |
|---|---|---|
| 1. | Dr. Nazmi Almasri | TESOL – IUG |
| 2. | Prof. Dr. Ezzo Afanah | Doctor of curriculum studies – IUG |
| 3. | Dr Akram Habib | English Literature - IUG |
| 4. | Dr. Sanaa Abu Dagga | Doctor of Researches Evaluation– IUG |
| 5. | Dr. Jamil Nashwan | Doctor of curriculum studies – Al Aqsa University |
| 6. | Mr. Alaa Harb | Supervisor of English – UNRWA |
| 7. | Dr. Kamal Murtaja | TESOL – IUG |
| 8. | Dr. Ahmed El Nakhalah | TESOL- Al Qds Open University |
| 9. | Dr. Ibrahim El Masharwi | TESOL- Al Qds Open University |
| 10. | Dr. Sohail Diab | TESOL- Al Qds Open University |

**Palestinian National Authority**

Ministry of Education & Higher Education

Deputy Minister Office

السلطة الوطنية الفلسطينية

وزارة التربية والتعليم العالي

مكتب الوكيل

الرقم : و ت غ / مذكرة داخلية ١٣٤٥

التاريخ : 2007/5/23

السادة / مديرو التربية والتعليم – محافظات غزة      حفظهم الله ،،،

السلام عليكم ورحمة الله وبركاته ،،،،

الموضوع : تسهيل مهمة بحث

يقوم الباحث/ حسني محمد صبح، والمسجل لدرجة الماجستير في أصول التربية

تخصص مناهج وطرق تدريس اللغة الإنجليزية بالجامعة الإسلامية، بعمل بحث بعنوان

Evaluating the Assessment of 10<sup>th</sup> Grade English Writing at Gaza

Schools in the light of the modern Trends"

لامانع من قيام الباحث من تطبيق أداة بحثه وهي استبانة على عينة عشوائية من معلمي اللغة

الإنجليزية للصف العاشر، وذلك حسب الأصول .

وتفضلوا بقبول فائق الاحترام ،،،

د. محمد أبو شقير

وكيل وزارة التربية والتعليم العالي

نسخة : الملف

140

# Appendix II

## Alternative assessment Questionnaire

The researcher is carrying out an MA research on ***Evaluating English Writing Assessment in the 10th Grade at Gaza Schools with Regards to the Contemporary Trends***

Based on modern trends of assessment found in the literature, the researcher classified the most common used methods of alternative assessment. They include:

1-assessment    2- observation    3-self- assessment    4- peer-assessment    5-journals 6-logs  7-anecdotal records  8-conferences and interviews   9-portfolios   10-rubrics

You are kindly invited to complete this questionnaire about the methods you use to assess your students' writing. Your response will provide valuable insight into the assessment methods used in Gaza schools. Please answer each question in the following  parts to the best of your ability.

**The key for your answers is:**

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| **Never** | **Rarely** | **sometimes** | **Often** | **Very often** |

 **Please indicate your gender**                    **Years of experience**
**Female** 
**Male**                                               **……………  years**

**1-Testing**

**A: If you use the following types of tests to assess the 10<sup>th</sup> grade students' writing during the academic year, please indicate which of the following you use.**

| No | Items | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 1 | Objective tests e.g. true/false, multiple choice, multiple-response and matching questions | | | | | |
| 2 | Subjective tests e.g. extended-response questions and essays. | | | | | |
| 3 | Diagnostic test | | | | | |
| 4 | Quizzes | | | | | |
| 5 | Others(please specify):  1……………………………………………… 2………………………………………… 3………………………………………… | | | | | |

**B:  If you do not use the previous methods, indicate the reasons.**
1    They may not sample a wide range of content
2    Subjectivity may affect fair grading.
3    Others…………………………………………….
---------------------------------------------------------------------------------------------------------

**2. Portfolio**

A portfolio is the collection of samples of students; work during the academic year
**A: If you use portfolios to assess the 10<sup>th</sup> grade students' writing during the academic year, please indicate what you collect of the following:**

| No | Items | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 1 | Students' best written work. | | | | | |
| 2 | students reflection on why they chose the pieces they did. | | | | | |
| 3 | samples of  writing tests and quizzes | | | | | |
| 4 | self- assessment checklists filled by students | | | | | |
| 5 | Others(please specify):  1……………………………………………… 2………………………………………… 3………………………………………… | | | | | |

**B:  If you do not use the previous methods, indicate the reasons.**
1    They may increase my work load.
2    They are time- consuming.
3    Others………………………………………….

## 3 -Self assessment
**Self-assessment refers to the involvement of learners in making judgment about their own learning .**
**A: If you use self assessment as a method to assess the 10<sup>th</sup> grade students' writing during the academic year, please indicate what you implement of the following:**

| No | Items | 1 | 2 | 3 | 4 | 5 |
|----|-------|---|---|---|---|---|
| 1 | Providing students with criteria check sheets (or have the class generate them) that specify exactly what constitutes a good product | | | | | |
| 2 | Training students gradually to assess their own work. | | | | | |
| 3 | Training students to identify specific areas where they need more support and can seek help. | | | | | |
| 4 | Providing examples of other students' (anonymous) work to discuss in terms of its strengths and weaknesses. | | | | | |
| 5 | Providing them with a description to the levels of quality of their work from excellent to poor. | | | | | |
| 6 | Others(please specify): 1…………………………………………… 2…………………………………………. 3………………………………………. | | | | | |

**B: If you do not use the previous methods, indicate the reasons.**
1 They may increase my work load.
2 They are time- consuming.
3 Others…………………………………………….

-------------------------------------------------------------------------------------------------------
## 4-Peer assessment
**Peer assessment is the process whereby groups of individuals rate their pairs**
**A: If you make students assess each other's writing during the academic year, indicate what you implement of the following:**

| No | Items | 1 | 2 | 3 | 4 | 5 |
|----|-------|---|---|---|---|---|
| 1 | Training students over time to assess the work of other students | | | | | |
| 2 | Providing students with basic assessment criteria by which they assess their peers. | | | | | |
| 3 | Using examples of work from anonymous learners as models of peer assessment. Encourage learners to. | | | | | |
| 4 | Providing students with opportunities to assess their peers in situations in which their assessments do not affect their grades. | | | | | |
| 5 | Encouraging student to share their and discuss what might improve each other's work | | | | | |
| 6 | Others(please specify): 1…………………………………………… 2…………………………………………. 3………………………………………. | | | | | |

**B: If you do not use the previous methods, indicate the reasons.**
1 Lack of support by supervisors, principals or others to apply this method.
2 I have not received any type of training to implement them.
3 Others………………………………………….

## 5- journal

**Journals refer to students' written record of a their expressions, experiences, feelings and reflections which could be in a notebook or a file**

**A: If you assess students using journals, indicate what of these items the journal include**

| No | Items | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 1 | Document of their expressions which the use when they write | | | | | |
| 2 | Document of their feelings and reflections about the writing lessons and tasks | | | | | |
| 3 | Document of their experiences and progress. | | | | | |
| 4 | Some of the difficulties they face in doing the task and how interesting or useful they find the tasks they did . | | | | | |
| 5 | Others(please specify):  1……………………………………………… 2…………………………………………. 3…………………………………………. | | | | | |

**B:  If you do not use the previous methods, indicate the reasons.**
1   It can be difficult to determine assessment criteria.
2   They may increase my work load.
3   Others………………………………………….

---

## 6-interviews and Conferences

If you conduct interviews and conferences  to assess students' writing during the academic year, indicate what you do of the following**.**

| No | Items | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 1 | Identifying what students have learned and what they have yet to learn. | | | | | |
| 2 | Assessing instructional changes or innovations | | | | | |
| 3 | Giving individual feedback and clarify misconception | | | | | |
| 4 | Exploring individual differences in experiences and outcomes | | | | | |
| 5 | Recording information during or immediately following the communication | | | | | |
| 6 | Others(please specify):  1……………………………………………… 2…………………………………………. 3…………………………………………. | | | | | |

**B:  If you do not use the previous methods, indicate the reasons.**
1   It can be difficult to determine assessment criteria.
2   They may increase my work load.
3   Others………………………………………….

**7-Rubrics**

**A rubric is lists of criteria handed out before the assignment begins in order to get students think about the criteria on which their work will be judge**

**A: If you rubrics as a method to assess students' writing during the academic year, indicate what you do of the following.**

| No | Items | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 1 | Providing students with transparent criteria against which to judge their own work. | | | | | |
| 2 | Providing students with a description to the levels of quality from excellent to poor. | | | | | |
| 3 | Sharing the descriptions of the scale with the students and ask for feedback so that each level is clearly understood by students | | | | | |
| 4 | Providing examples of students work that illustrate each performance level. | | | | | |
| 5 | Giving students feedback clearly related to the assessment criteria. | | | | | |
| 6 | Others(please specify): 1…………………………………………… 2…………………………………………. 3………………………………………… | | | | | |

**B: If you do not use the previous methods, indicate the reasons.**
1 It can be difficult to determine assessment criteria.
2 They may increase my work load.
3 Others……………………………………………

---

**8- observation**

**A: If you use observation as a method to assess students' writing during the academic year, indicate which of the following you do.**

| No | Items | 1 | | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 1 | Observing students' attendance | | | | | |
| 2 | Observing behavior | | | | | |
| 3 | Observing performances | | | | | |
| 4 | Using different tools to observe students | | | | | |
| 5 | Writing short notes during a lesson, as students either work in groups or individually, or after a lesson. | | | | | |
| 6 | Others(please specify): 1…………………………………………… 2…………………………………………. 3………………………………………… | | | | | |

**B: If you do not use the previous methods, indicate the reasons.**
1 It can be difficult to determine assessment criteria.
2 They may increase my work load.
3 Others……………………………………………

# 9-Anecdotal records

An anecdotal record is a collection of brief, written observations of students related to their progress in learning using note cards or a notebook.

**A: If you use anecdotal record to assess students' writing during the academic year, indicate which of the following you use.**

| No | Items | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 1 | Using anecdotal records to assess students participation and effort during the writing tasks | | | | | |
| 2 | Using anecdotal records to trace students' gradual progress towards attaining identified writing objectives | | | | | |
| 3 | Practicing this method systematically after each writing lesson or at the end of the day | | | | | |
| | Others(please specify):  1………………………………………………… 2…………………………………………………. 3…………………………………………….. | | | | | |

**B:  If you do not use the previous methods, indicate the reasons.**
1    They are time- consuming.
2    They are difficult to be applied in classless of large numbers of students
3    Others………………………………………………….

---

## 10-Checklists

**A checklist is a list of items to indicate if students can accomplish the listed writing objectives. Nothing is included about the quality of performance.**

**A: If you use checklists to assess students' writing during the academic year, indicate which of the following you use.**

| No | Items | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 1 | Using checklists to indicate students in self and peer assessment | | | | | |
| 2 | Using checklists to record observations during the writing task | | | | | |
| 3 | Others(please specify):  1………………………………………………… 2………………………………………………… 3…………………………………………….. | | | | | |

**B:  If you do not use the previous methods, indicate the reasons.**
1    I  have not been trained to use them.
2    They may increase my work load.
3    Others………………………………………………….