

The effect of using test based paper instead of copybook test to improve Iraqi educational test system by saving time and money

Dr. Hayder Ali Abdulridha AlZubaidi, Iraq

EFL at al-Muthana preparatory school/ Wasit

Master degree in ELT Methodology, evaluation and curriculums,

Hayder.alia@uokut.edu.iq

"The profoundly commendable test should be economy in time, effort, and expenses. A good test should not be expensive, very long, and tiring. Ease of Administration and Scoring. There should be full, clear directions so that the test administrator can perform the task quickly and efficiently."(Harris, 1969,p.p. 21)

" Good test should be economic . The cost in time, money and personnel of administrating a particular test."(Al-Juboury,1999. pp.29)

Introduction

In spite of the efforts to promote the Iraqi educational test system process and build economic test by trying out different teaching methods and/or techniques to instruct, there has been a growing concern the test system have not progressed as much as it was anticipated (AlZubaidi. 2014).

Iraq is pursuing a series of emergency measures to finance its 2015 budget, following years of poor governance and unchecked spending that led to a financial crisis now made worse by low oil prices and high military spending demands.

The extraordinary measures underscore the severity of Iraq's financial crisis, which was created by years of unsustainable budgeting practices and poor fiscal management. The order came from the Iraqi government to reduce spending as much as each ministry could . In fact one of the most spending lay in ministry of education in Iraq. The free education hold many issues is on Iraqi budgeting practices. So the current study was designed to identify the importance of removing test copybook from Iraqi testing system. This study also aimed to investigate how to make the test more economic and easy both to students and teachers by saving money and time. Three classes from Al Muthana preparatory school participated in the quantitative part of this study, while five instructors took part in the qualitative part. Based on the findings from the students' questionnaires, Al-Muthan students were found to be supported using based paper test instead of using test copybook.

The main themes that emerged from the interview data with the instructors and the students included instructors having clear ideas about what papers based test are, and they were really used it in monthly exam of English. All of them agreed that based paper test more easy to correct and saving time both to teachers and students. They believe it profoundly commendable to prevent students to cheat by using their paper .

Decisions related to the improvement of teachers practice constitute an important part of their work . Such decisions relate to a wide spectrum of issues , including assigning grades to students, evaluating suitability of

frequency in relevant literature.

textbooks and assigning students to an appropriate class. Moreover, it includes dealing with the designing classroom assessment and evaluation and improving their teaching. In order to carry out these tasks, teachers need more than access to different assessment techniques and instruments, they also need an understanding of the nature and purposes of evaluation, procedures for collecting and interpreting different kinds of information about students learning and their educational performance, the ability to make appropriate decisions about instruction and instructional plans that can have a significance impact on them (Marianne, 2001. p.p5-7).

One of the aims of EFL teaching process is to develop and conduct a reliable and a valid language assessment. Moreover, it would provide a wide range of both formal and informal evaluation techniques that can improve language teaching and learning processes (David & Ann, 2001:63).

EFL teachers are actively and continuously involved in language evaluation, sometimes as a person who makes actual decisions, sometimes as a collector of relevant information for others who will make the decisions, or as a helper of others in making decisions through offering interpretations on students performance (Richards, 1998).

Ramsden (1992) believes that evaluation of courses and programmes is a core activity and integral part to all teachers. In this sense, it is understood to be an aspect of the professionalism of teaching staff and indicative of a commitment to high quality personal standards of performance.

Evaluation is a way of understanding the effect of teaching on students' learning. It has three main techniques: gathering data about teaching activity, interpreting them and reflecting upon them in a critical manner. So, it is important to distinguish among these three components because they are distinct. Information about teaching and learning is rarely meaningful by itself. It becomes meaningful when it is interpreted. Meaningful interpretations are needed in order to decide what actions to take or what changes to make (Hatton & Smith, 1995).

Thus, the Problem of the study is best expressed through the following research questions:

- 1- Are Iraqi education test system economic to save time and money?
- 2- How can make the test system more economic?
- 3- What is the effect of removing the copybook test and replace it by based test paper?

To the best knowledge of the researcher, no study that answers the questions raised above has so far been conducted in Iraq. Such questions need to be answered empirically. Therefore, carrying out such a study is expected to contribute to a full understanding of the role of test based paper to save time and money. The test plays in improving Iraqi EFL teachers practice and thus, helps to fill the gaps of information in this area.

Aims

The study aims at:

1. Identifying Iraqi education test system
2. Finding out whether or not the effect of test based paper to save money and time.
3. Finding out the effect of test based paper on EFL students' and teachers' achievement.

frequency in relevant literature.

Limits

This study is limited to :

1- Fourth preparatory school teachers and students in Wassit Directory of education during the academic year 2014-2015.

Value

This study is hoped to be of value to :

- 1- The researchers and experts in ELT, since , this study presents a survey of the characteristics and techniques of test system.
- 2- EFL teachers and supervisors.
- 3- Iraqi government to help passing the financial crisis in Iraq now
- 4- Creating modern economic test by saving money and time
- 5- The study is the primary study to improve electronic test system in Iraq.

Assessment

It includes any means of checking what students can do with the language. It may be carried out before, during or after a course, or it may not even be connected with a course. It may be for individual students or for a whole class. Assessment is "one of the most potent forces influencing education" (Crooks, 1988, p.p.448). It has always been an integral part of FL pedagogy and education in general to the extent that learning and teaching are considered inseparable from assessment practices. Its crucial importance lies in the powerful impact it has on learning, teaching, curriculum and consequently on learners and teachers.

Defining Assessment

Assessment can be defined informally or formally. Informally it can be any appraisal, remark or judgment of a student's work or performance (Sadler, 1989, p.p.119). Formally it can be defined as "the process of defining, analyzing, interpreting, and using information to increase students' learning and development" (Erwin, 1991, pp.15). Lynch (2001, pp.358) sees assessment as the "systematic gathering of information for the purposes of making decisions or judgments about individuals". Assessment is also seen as systematic tools, techniques, and procedures used to collect, classify, and interpret information about students' capability or the quality or success of a language teaching and learning course according to different sources of the students' performance. A product or outcome of inferring about the students' knowledge is also referred to as an assessment (Nunan, 2001; Bachman, 2004; Linse, 2005).

Sometimes the terms evaluation and assessment are used interchangeably but they are not exactly the same. Nunan (2001) and Brindley (2003, as cited in Linse, 2005) draw a clear distinction between the two concepts. To them evaluation is the collection and interpretation of information about various aspects of the curriculum (including learners, teachers, materials, learning arrangements, etc.) for decision making purposes – the effectiveness of an education program; whereas assessment is a "subcomponent of evaluation". Assessment refers to the collection of information on what individuals and group of learners can do.

Testing is another term sometimes considered a synonym to assessment, yet they are not the same. Brown (2004) states the difference by saying that tests are prepared

frequency in relevant literature.

administrative procedures, usually administered within strict time limitations, to sample the performance of a test-taker in a specified domain; whereas assessment is an ongoing process that compasses wider meaning; assessment occurs all the time. Whenever a student responds to a question, offers a comment, or tries out a new word or structure, the teacher subconsciously makes an assessment of the student's performance. Tests, then, are a "subset of assessment", they are not the only form of assessment that a teacher can make (ibid). Assessment is larger in sense and includes both measurement tools which yield quantifiable scores and other tools of qualitative assessment, such as observation, journals, portfolio, etc (Davies et al., 1999).

In language teaching and learning, Brown et al. (1994) remind us that it is important as providers of education to step back and consider why we assess. They list some commonly seen reasons why assessment is useful. It: enhances motivation; creates learning activities; gives feedback to the student to help him identify the strengths and weaknesses; gives feedback to the teachers on how well the message is getting across; judges performance (grade/degree classification); and provides quality assurance-internal and external to educational institutions and others.

Students might be tested both at the beginning and at the end of a course to assess the quality of teaching in the course. They may also be tested at both sides of the teaching process to assess their level of achievement. Assessment can provide inferences about a person's potential, intelligence, aptitudes, attitudes, motivations, etc. and then the worth is given in the form of grades, then recommendations are made.

Most forms of assessment have formative or summative functions. Each form has its own techniques. Summative assessment activities are used to judge or sum up achievement. Formative assessment activities can be used to improve performance or practices. Formative assessment encompasses any activities undertaken by teachers or students that provide feedback that will be used to modify teaching and learning activities (Black & William, 1998a).

Purposes of Assessment:

- 1- Placement- to place or classify students on a suitable course.
- 2- Diagnosis- to specify the student's particular strengths & weaknesses..
- 3- Selection - to select fairly the best candidates for limited number of places or jobs or for higher education.

- 4- Evaluation - to evaluate the effectiveness of the syllabus as well as the method of teaching so as to makes adjustments when needed.

- 5- Progress - to assess what progress they have made.
- 6- Prediction - to discover potential abilities and aptitudes and to predict probable future successes whether in school or outside.

- 7- Accreditation – to provide evidences that the student has undergone a course of training or studies and is qualified to practice his profession.

Types of Assessment

frequency in relevant literature.

Assessment is used for different purposes at different points in the learning process. The following are four key terms which appear with some frequency in relevant literature.

Achievement / Attainment tests: they are designed to measure the extent to which the students have mastered the material taught in the classroom. They aim at stimulating learning and reinforcing what has been taught. Annual school examinations and public examination are far more formal than class progress tests and measure achievement on a larger scale.

Proficiency test:

they measure what tests have learned but they are not related to any syllabus or teaching program, i.e. they aim to assess global ability. They are used in placing individuals in learning groups appropriate to their level of knowledge.

Aptitude tests:

they measure the student's probable performance in a foreign language he has not started to learn. Language learning aptitude consists of factors such as intelligence, motivation, age, memory, etc.

Diagnostic tests:

they aim to find out the strong and weak points of the class as whole and individual students so that appropriate remedial action can be taken.

Modes of assessment:

Formal assessment :- Is standardized large scale tests which are prepared by professional testing services to assist institution in the selection placement and evaluation of students.

Informal assessment :- Is often used diagnostically by the teacher to obtain reliable insights about the students ability and state of development.

Continuous assessment :- means continuous updating of judgments about students performance in relation to specific subject

Terminal assessment :- means assessment that takes place at the end of the course

Internal assessment :- Is usually carried out by the teacher of the subject once a month or a year to check students progress .**External assessment** :- Is usually constructed by a

team of specialists or an organization outside the school and applied on thousands of students for the purpose of selection or placement .**Examinations** :- are either school examinations constructed by the teacher or teachers of the subject in a given school , or Ministry examinations constructed and administered by the Ministry of Education

Coursework :- Is work completed by the students under the supervision of their teachers . such as (assignment or fieldwork , which is a piece of work which has a unity of theme or objective .

Formative Assessment

It helps to show which aspects the student has mastered and where remedial work is necessary .

frequency in relevant literature.

Formative assessment is a method designed to establish how much progress a student is making during learning with a view to give feedback to the student (Atkins et al., 1993). An example is the essay used as graded coursework. Comments written on it by the teacher are intended to help the students learn from mistakes and improve their work (Gibbs & Habeshaw, 1990). Formative assessment is often referred to as assessment of the *process* of learning, and is primarily used to improve learning by giving the student information on his or her learning progress while still learning. Formative assessment can be given either by one-way communication from the teacher to the student, or in conference with the student. The major difference between the two, according to Gipps (1994), is their purpose and effect. Major arguments for developing formative assessment practices are democratic in essence, that is to promote and improve learning for all students, and leading to empowerment and self regulation.

Formative assessment includes all activities that provide information that is used as feedback to adapt teaching and learning in the classroom to student needs, and to promote student learning (Black, 1998; Black & William, 1998b; Black et al., 2003, 2004). Spolsky (1992) rightly argues that diagnostic or formative assessment is typically curriculum-driven. In an extensive survey of the research literature, Black and William (1998b) and Black et al. (2003) concluded that formative assessment raised standards. Second, there was evidence that it helped low achievers more than other students, reducing the range while raising achievement overall. Third, frequent assessment feedback helped both groups, lower and higher achievers, enhance learning. Fourth, formative assessment can be a powerful weapon to create a "culture of success" in the classroom, negating low self-esteem, low self-efficacy and inadequate learning approaches, so that all students are able to achieve (Black & William, 1998b; Black et al., 2003). Feedback should be about the students' work, in relation to previous performance and set criteria, not about the self or amount to a comparison with others. It should consist of concrete advice on how to improve.

Summative Assessment

The summative assessment :- Is usually given at the end of a course to measure the sum total of the material covered.

This assessment method is designed to establish what a student has achieved at the end of a unit or course. A final mark or grade is awarded (Atkins et al., 1993). In contrast with formative assessment, i.e., the assessment of *process*, summative assessment equates with the assessment of the *product* or *outcomes* of learning.

Summative assessment is generally meant to measure the achievement of learners over a given period of time, such as at the end of a course module, and it is also used for standardized tests where a grade or score need to be given in order to improve a learner's ability. Scores and grades from these tests usually have high stakes for those involved, and used to compare the quality of schools, language programs, job applicants, quality of university, and a wide range of other categories. For summative assessment, " the focus of attention is on the result of learning [and is] concerned with identifying overall levels of achievement and measuring what learners do against them"(Rea-Dickens, 2000). This kind of assessment is usually called 'assessment of learning', and is usually not intended to guide the learning process in any way, but merely to measure the learning that has taken place.

frequency in relevant literature.

Often the stakes of summative assessment, especially from external assessing sources, are incredibly high and have a huge impact on our lives and futures. The qualifications gained from testing can often define our whole character to others around us, and act as gateways at important transitional moments in our education and employment. One would think then, that high stakes tests would provide all the motivation for the learner that is necessary for learning to take place. However, according to studies carried out by McNeil and Valenzuela (1998), reviewed in Harlen and Crick (2003), high stakes tests "do not have a markedly positive effect on teaching and learning in the classroom. High stakes tests do not motivate the unmotivated".

There is also a huge amount of pressure on teachers, and students to pass these tests, a trend reviewed in Harlen and Crick (2000) who find out that high stakes tests increase test anxiety, increase the pressure on students to do well, focus teaching on the content of the tests and lead to students adopting test-taking strategies that are harmful to learning development.

A major component of summative assessment is grading. Broadfoot et al. (1992, p.p.6) mention that in terms of assessment which leads to labeling of student's abilities with a grade, usually denoted by the letters A to F, or on a scale of 0-10, this form of assessment can "in certain circumstances have harmful effects on children's learning ". They mention further:

There is much research ... that shows that a competitive climate, such as labeling produces, does not increase learning outcomes for all. Rather the already successful thrive and the less successful (always the majority) underachieve more and more as they are repeatedly discouraged by the way their efforts are judged.

This view is supported by Torrance (2007, p.p. 287), who claims that achievements in summative assessment are "not necessarily the highest grades available or even directly related to competent practice [...] provided the work is (eventually) completed, the award will be achieved". More dangerously in terms of summative assessment and its effects on teaching and learning process, Brown (2005) adds the following negative effects summarized into four categories as follows:

1. Teaching factors include:
 - Teachers narrow the curriculum; teachers stop teaching new material and turn to reviewing material; and teachers replace class textbooks with worksheets identical to previous years' tests;
2. Course content factors:
 - Students practicing "test-like" items similar in format to those on the test; students applying test-taking strategies in class; and; students studying vocabulary and grammar rules to the exclusion of other aspects of language.
3. Course characteristic factors:
 - Students being taught inappropriate language-learning and language-using strategies, reduced emphasis on skills that require complex thinking or problem-solving, courses that raise examination scores without providing students the English they need in language interaction are also called test score 'pollution'; and a tense atmosphere in the class.
4. Course time factors:

frequency in relevant literature.

Enrolling in, requesting, or demanding additional (unscheduled) test-preparation classes or tutorials (in addition to or in place of other language classes), review sessions added to regular class hours, skipping language classes to study for the test, and lost instructional time.

What is clear from what has been mentioned above is that summative assessment may have negative washback on teaching and learning process. Teachers are either consciously or unconsciously directing their teaching to help students pass their tests and get good grades at the expense of other areas of learning, and the students are focusing too closely on the processes of passing these tests, perhaps at the expense of their personal development or going against their natural learning style. In fact, what has been mentioned above characterizes the Iraqi EFL situation.

As a result of all the above-mentioned negative impact of summative assessment on teaching and learning, the researcher will enhance the use of formative assessment in his experiment.

What is a Test?

Test is a method of measuring a person's ability or knowledge in a given area. There is a set of techniques, procedures, test items, that constitute an instrument of some sort and that method generally requires some performance or activity on the part of either the testee or the tester or both (Brown, 2004).

Heaton (1987) states that the test which measures the students' performance, as fairly as possible, can be used to motivate the students and a well-constructed test will provide the students with the opportunity to show their abilities to recognize and produce correct forms of the language.

Language Testing: A Historical Perspective

Madsen (1983, p.p. 21) indicates that "language testing has evolved through three major stages, which reflect people's attitudes towards the goal of language teaching and language learning", which are the intuitive, scientific and communicative stages. While Savignon (1985) says that the selection of teaching materials is of course inextricably tied to testing. The recent history of second-language testing, which is called the "psychometric-structuralist period, is characterized by a widespread interest in the use of psychological measurement techniques to evaluate learners' mastery of discrete structural features of language" (Spolsky, 1978, p.p. 22).

Power (2004, p.p. 6) raises the question "How do we design a test that mirrors the students' abilities?" To answer this question, it is important to design a test that reflects what had been taught (Nunan, 1988). The test should be seen as working with the students not against them; it should offer a constructive help rather than non-directed criticism; it should show them how they were faring on the road to fluency and point the way forward; and it should be about assessment of individual abilities and

frequency in relevant literature.

improvement rather than a comparison of proficiency levels. The success of this test would be judged primarily by its effectiveness in favourably affecting the student's perception of his/her language abilities (Nantle -Bromley, 1995).

Test Making and Test Taking

Test Making :The Construction of Test Items

Gronlund (1981, p.p. 155) states that:

The construction of good test items is an art. The skills it requires, however, are the same as those found in effective teaching. Needed are a thorough grasp of subject matter, a clear conception of the desired learning outcomes, a psychological understanding of pupils, sound judgment and a touch of creativity.

For the purpose of this value, the constructions of test items will be regarded as consisting of the following steps:

1. Planning the test which encompasses all of the varied operations that go into producing a test such as paying attention to the content of the test items, types of test items, directions to the testees, etc. (Lindquist,1999)
2. Writing the test items which is an art, and is essentially creative. Each item presents new problems and new opportunities.
3. Trying out the test in preliminary form and assembling the finished test after trying out. The test items must be tried out on a sample of testees. This sample should be as nearly like the population with whom the final form of the test is to be used.
4. Determining the procedures in scoring of tests items which include the scoring formula, the weight of items, the types of keys to be used, etc.

On the basis of this issue, Horrocks and Schoonover (1968) mentions that tests are either made by the teacher (teacher made tests) or by a group of experts (standardized tests) and these types of tests differ in a number of ways.

Teacher- made tests are constructed by the teacher for use in a specific situation while the contents of standardized tests are generally applicable for any one at the age or grade level at which they are intended to be used. Directions are uniform, and, time is specified, and the scorer provides a key to the answers .All this insures that each student taking the test takes it under “standardized” condition.

Test Taking

Test taking needs every effort to be made so as to enable students to approach a test with a positive attitude and with recognition that the test supplies an opportunity for self- assessment and self-analysis. If the students view the test as a learning aid, much of the emotionality connected with test-making will be removed. In the Write Source (2000), there is a brief presentation of this skill with some practical procedures to be followed by any testee. In the fore- mentioned reference book it is stated that the key to doing well in any academic study is simple-be involved and be prepared. This is never more true than when a student is about to be tested on what he/she has learned. Consequently there are several skills or guidelines which a student should follow to help

frequency in relevant literature.

him/her as to begin his/her preparation. The necessary test-taking skills are:

1. Organizing and preparing test material.

This step is achieved through the following:

asking questions about the material, organizing the taken notes, and thinking about the test by gathering old quizzes and exams, preparing an outline of everything to be covered on the test, and then set aside a list of any question one needs to ask the teacher or another student.

2. Reviewing and remembering test material. This is achieved by the following steps: beginning to study early, using study aids such as lists, diagrams, flash cards, note cards, etc., and testing oneself by trying to imagine what questions will be on the test and then write answers to them. Taking the test: This phase involves three major steps: before beginning phase: it is achieved by checking to see that all the materials (paper, pencil, books, etc.) needed to take the test are available. during the test phase: This step is conducted through the following sub-steps: Skimming the entire test quickly, estimating how much time needed for each question, then beginning the test and watching the time carefully, Reading the directions carefully, Reading all questions carefully, paying attention to words like only, all, and never, and answering the questions one is sure of first, then moving to another question when one is really stuck. Before finishing phase: This is done through the following: double-checking each section or page to make sure of doing all the questions and checking the handwriting thoroughly, and asking the teacher about any confusing question to be sure of the answer.

To sum up, two factors are related to test taking which take place during and after the test:

The relationship between what is studied and what students think they will be tested on is necessary to be recognized. Students will learn what the teacher asks about and the level that is required rather than what is presented (Horrocks, 1999; Valette, 1999).

Characteristics of a Good Test

Most methodologists, if not all, such as (Andrew, 1980: 133-40; Grant, 1987; Harmer, 2001, and Power, 2004) generally believe that a good test is based on the ground of the following criteria:

1. **Validity:** A test is valid if it tests what it is supposed to test (Harmer, 2001). A question may be raised: How does one establish the validity of a test?

Statistical correlation with other related measures is a standard method. Lado (1969, p.p. 169), for instance, mentions that “validity is often expressed as a coefficient of correlation between the test and the valid criterion”. But ultimately validity can only be established by observation and theoretical justification. There is no final, absolute, objective measure of validity. The test should accurately and sufficiently measure the testees for the particular purpose, or objectives or criteria of the test (Brown, 2004:). The three types of validity which are considered important. As Taeduck and Finch (1998) list them, are as follows:

1. Face validity.

frequency in relevant literature.

2. Construct validity.
3. Content validity.

2. Reliability: Davies, et al. (1999, p.p. 168) state that:

Test reliability is the actual level of agreement between the results of one test with itself or with another test. Such agreement ideally would be the same if there were no measurement error, which may arise from bias of item selection, from bias due to time of test or from examiner bias.

It is worth mentioning that both validity and reliability are going to be defined and discussed thoroughly in Chapter Four.

3. Practicality: Another characteristic of a good test is its practicality or usability, i.e., the test is beyond the tester's means or facilities (Harris, 1969;,p.p. 21). He expresses that a number of practical considerations must be kept in mind in the preparation of a new test which are:

1. Economy

There should be economy in time, effort, and expenses. A good test should not be expensive, very long, and tiring.

2. Ease of Administration and Scoring

There should be full, clear directions so that the test administrator can perform the task quickly and efficiently.

2.6 The Principal Educational Uses of Language Tests

In its broadest sense. Language tests have many uses in educational programmes, and quite often the same test will be used for two or more related purposes. They are used to:

determine readiness for instructional programs, classify or place individuals in appropriate language classes, diagnose the individual's specific strengths and weaknesses, measure aptitude for learning, measure the extent of students' achievement of the instructional goals, and evaluate the effectiveness of instruction.

2. Subjective Tests

Subjective tests (henceforth Subj. tests) are any scoring procedure which involves the exercise of judgment by the scorer (Davies,et al. 1999).They may result in a range of possible answers, some of which are more acceptable than others (Power, 2004).

Subj. tests give the test taker fewer cues and open wider areas for responses

frequency in relevant literature.

.Subj. tests require students to organize their own answer, expressed in their own words and this improve their writing ability (ibid.)

2.7.1 Advantages of Subjective Tests

Advantages of subj. tests are listed and briefly explained below:

Subj. tests provide a freedom of response that facilitates assessment and evaluation.(McIntire and Miller,2000:204). They help the learner to focus on demonstrating what they have learned and do not limit them to specific areas. They are easier to construct and revise. They measure students' ability to use the language correctly in their own writing (Lindquist, 1974). They measure the various aspects of students' progress toward the acquisition of writing (Valette, 1977). They measure learning outcomes that cannot be measured by objective test items such as the ability to select, organize and evaluate ideas.

They can be presented on chalkboard. They can give a piece of evidence of depth of comprehension. They offer a chance to describe relevant applications (Horrocks & Schoonover, 1968).

Disadvantages of Subjective Tests

The following disadvantages of subj. tests are considered:

Subj. tests require lengthy written responses by the test takers (Gronlund, 1981). The scorer judgments can be influenced by writing skills ranging from legibility to graceful phrasing and such considerations can lead to inaccurate test scores.

It is up to the tester to decide what is a good or poor answer because students may give a variety of responses each some -what different from the others (Valette, 1977). Subj. tests are more difficult; they require independent scoring by two experts to increase reliability and accuracy (Power, 2004). The tests are time-consuming to score. The tests provide a limited sampling of information. Reliability is low in subj. tests. The verbally facile students may depend upon quantity of production not quality. Validity may be low because of the limited sampling and reliability. The testees require a long time to respond only to a few items. The questions are not sufficiently specific and encourage glittering generalities. (Horrocks & Schoonover, 1968).

Scoring Subjective Tests

On the theme of scoring subj. tests, the following points are taken into consideration:\The scorer awards points based on how closely the test takers' response matches a predetermined correct response (McIntire & Miller, 2000). The scorer can use the “global quality sealing” by putting in rank order after a quick scan (Power, 2004). The scorer can assess in categories (e.g. vocabulary, grammar, content and form) 5 marks to each; total mark out of 20 . The scorer can divide the answer into sense groups, i.e., the marking system can be based on communication and correctness as two separate criteria (Valette, 1999).

Types of Subjective Tests

frequency in relevant literature.

Subj. tests are categorized into the following: Transformation items: students are asked to transform sentences or passages according to a specific pattern. Directed questions. Translation. Controlled or guided composition and free composition (Valette, 1999).

Techniques of Written Subjective Tests

Essay Tests

These tests are popular subj. tests. They are usually general in scope and require lengthy written responses by test takers. These kinds of tests provide a freedom of response, i.e., the responses generated can vary in terms of width and depth of topic (McIntire & Miller, 2001).

Sentence Completion

The students are required to express the best way of completing a sentence in terms of grammar, dictation, tone, sense and information. Harris (1969, p.p. 74) indicates that "This item type provides an excellent method of measuring a wide range of problems relating to the effective use of written English".

Completion tests consist of an incomplete piece of writing, and the tester asks the students to complete it. This kind of test has three major weaknesses. First, it is the difficulty of preparing a context that will allow only one possible answer. Second, the time required for scoring. Finally, the difficulties of spelling that the students may face when writing (power, 2004).

Composition Test

Davies et al. (1999) comment that "composition tests are common in tests of English for academic purposes, where the purpose of the test is to predict which candidate will best be able to cope with the demands of academic essay writing". Lado (1962, p.p. 250) states that the virtues claimed for this technique are the realistic nature of the response and the fact that it can show how well a student can think in the language, while the shortcomings of this technique are the difficulty of scoring the responses objectively.

Restricted Response Questions

The restricted response questions tend to limit both the content and the form of student's response. The content is usually limited by restricting the scope of the topic to be discussed. On the other hand, the form of responses is restricted through the statement of question. (Gronlund, 1981). There are several types of restricted questions such as:

Answer briefly [summarize the contents of ...]. Describe with examples [analyze and apply]. Why do you think [explain or justify]. List [create as many ways as you can think of....]. Define [state the main points included in

frequency in relevant literature.

Extended Response Questions

As a matter of fact, these questions provide a very wide range of latitude, i.e., the students are free to select any information that they think is important, to organize the form of the answer to evaluate and to integrate the ideas. There are several types of extended response questions such as:

Compare (Describe the similarities and differences). Write a scientific evaluation [evaluate and relate cause and effect].

Objective Tests

Generally speaking, educators look at objective tests (henceforth obj. tests) from the marking point of view ,i.e., it has only one correct answer. Davies et al. (1999, p.p.129) say that:

An objective item call for the choice of a correct answer out of several provided for each question. or, one clearly specifiable correct response-objective items typically contain a stem, which presents the problem clearly and then two or more choices. Only one of which is correct.

It is worth noting that it is not the test which is obj. but the system by which it is marked. An obj. test is made up of items each of which is discrete in itself, but shown to contribute to the total performance of the test.

Obj. tests tend to point towards measures of receptive learning, i.e., learners do not have to produce language, in many of the tests items (Power, 2004).

Advantages of Objective Tests

Actual marking is easy. Marking can be done mechanically or by over lay (Heaton, 1987). The test can be pretested. Large population can be tested. The tests can be compared over different years (Gronlund, 1981).

A good deal of judgment has been used in developing setting and accepting the tests items. A much wider sample of grammar, vocabulary and phonology can generally be included in an objective test (Davies et al.,1999). Test takers spend less time and effort in answering the items of the test (Power, 2004).

Disadvantages of Objective Tests

The following points are regarded as the disadvantages of obj. tests: They invite guessing (Davies et al., 1999). They cannot test students ability to communicate in the target language .There is no production of language. They cannot evaluate actual performance. (McIntire & Miller, 2000). Obj. test items are limited to measure simple learning out-comes, the main exception to this seem to be in distinguishing between fact and opinion. An obj. test may have a variety of answers of varying degree of correctness. Obj. test items are restricted to the measurement of factual information based on rote learning. It is highly susceptible to the presence of irrelevant clues. (Gronlund,1981).

frequency in relevant literature.

Scoring Objective Tests

The following points are born in mind when scoring obj. tests: In an obj. test correct answers are clearly specified. Test makers do not have to make judgments. A list of keys gives the only correct answer (Davies et al., 1999)

Types of Objective Tests

Obj. tests are classified into the following:

True/ False, Filling-in-the Blanks, Completion, Short answer, Alternative answer, Multiple-choice (which is used to measure the students' ability to identify the facts), Matching, Regrouping, Identifying , Multiple hot spots:(moving label to appropriate place on an image). Sequence(placing items in to a particular sequence)(Gronlund,1981)

Techniques of Written Objective Test

True/False

The true/false test is one of the most widely used tests. It is of considerable use in assessing the student's progress chiefly because they can be constructed easily and quickly.

Davies et al. (1999, p.p. 215) state that a true/false item is "A dichotomously scored test item, where only two options about a statement ("true" or "false") are offered to the test taker, although objectively marked, it has the disadvantage that a guess gives the candidate a 50% chance of answering the item correctly".

In constructing a true/false item, the following points should be considered: The stem: It usually asks "Which of the following is True?" The statements: They follow the stem and they can be either true or false statements. The true-false question has only two options. The odds are always 50-50 with this type of items. In order for a statement to be true, it must be so 100% of the time i.e. each part of the question must evaluate the trueness of how, what, where, when, and who for each statement. Words like (never, always, usually, not usually and some) give specific meanings and make confusion to the students (Mallory, 2002).

Fill-In the Blanks

Such tests differ from similar completion tests because sufficient clues are provided both in the sentence and in the list of words provided in the brackets and "These clues enable the student to know exactly which word is required" (Heaton, 1983:, p.p. 143). This type of tests consists of a sentence, from which a word(s) has been removed, and a group of words may be given before the sentences that have the missing words, the testee has to supply the most appropriate word(s) to fill in the blanks (Al-Jarah, 1991).

Underline

frequency in relevant literature.

This test is another type of obj. tests in which the testee is required to recognize certain sound(s) of a word, i.e., these tests require the ability to discriminate between English phonemes.

Although basically the testing of sound perception and recognition is not difficult in scoring, it needs instructions for the testee and requires certain standards of form, validation, etc. that call for additional consideration when preparing the test (Lado, 1962).

Matching

Generally speaking, this test is the simplest form. It consists of two lists of words, phrases, symbols, facts, etc. The students should match between the items to produce a correct answer. Alderson, et al., (1999 p.p. 51) state that "By "matching" we mean items where students are given a list of possible answers which they have to match with some other lists of words, phrases, sentences, paragraphs, or visual clues".

Obviously, the disadvantage of this item is that "once three of the four items have been accurately matched the fourth pair is correct by default". To avoid this, more alternatives should be given to the student. It is worth noting that in matching tests the tester must be sure that each item in the first column only matches one item in the second column (Power, 2004).

Re-grouping

In a rearranging test the students are asked to put a group of words, phrases, sentences, or paragraphs in order. The test items are difficult to construct, i.e., it is not easy to provide words or phrases which only make sense in one order (Alderson et al., 1999).

A Comparison between Subjective and Objective Tests

The issue of comparing subj. and obj. tests highlights the major strengths and weaknesses associated with these types of tests as shown below: Sampling: "objective tests provide ample opportunity to sample the test universe" (McIntire & Miller, 2000). The test maker can cover a very wide array of the subject matter and that increases the content validity of the test, while subj. test maker is limited to the number of items or questions to which the test taker can respond in one session and this leads to the weakness of content validity.

Difficulty: subj. tests are more difficult to control than obj. tests. The scorer judgments depends on some strategies, thus the results depend on the same strategies (Morton, 1991). Construction: obj. tests require extensive thought, development and time to balance responses in terms of content depth, length and appeal to the test takers .On the other hand, subj. tests need fewer items that are easier to construct and revise (Power, 2004). Response sets: Test takers can guess the correct answer in obj. tests or they can choose socially desirable responses. But in subj. tests the responses might bluff or pad answers with superfluous or excessive information (McIntire & Miller, 2000). Test developers prefer subj. to obj. tests on some occasions ,because subj. tests give the test taker fewer cues and open wider areas for response while obj. tests nly on cues provided by the test and produce no language at all(power,2004).

frequency in relevant literature.

Testing Language Elements

The components of language can be isolated to the following elements: Phonology (concerned with pronunciation, stress and intonation). Vocabulary (concerned with word meanings and word arrangements), and Grammar.

(Lefrancois, 1999)

Heaton (1987) mentions that test items that are designed to test phonology should assess at the same time the ability to: Recognize and pronounce the different sounds of the language, Recognize and use the stress patterns of a language, and Hear and produce the melody of a language (i.e., the rise and fall of the voice).

Although the sound system of a language is the same for the speaker and the hearer, Lado (1962) states that “the techniques used to test recognition...will differ from those used to measure production of the sounds of the foreign language”. Moreover, speaking is just as writing (a complex skill) which requires the simultaneous use of a number of different abilities which often develop at different rates(ibid).

Techniques in Testing Pronunciation

Pronunciation is tested globally by different techniques, which are either oral or written as the interview, conversational exchange, reading aloud oral drills, using pictures for oral productions, the short talk, and different types of written tests to provide students with an understanding of how pronunciation contributes to communication (Rogerson & Gilbert, 1990).

Content of tests:

Listening comprehension: including phoneme discrimination, stress and intonation recorded texts following by questions of varying techniques. *Speaking ability:* in the form of an interview, picture-description or role-playing. *Reading comprehension:* questions are given to test student's ability to understand a text by applying certain cognitive skills. *Writing ability:* usually in the form of the foregoing skills are: grammar, vocabulary and phonology.

Assessment can be introduced at different stages of a course of study - Terminal of final simply means assessment that takes place at the end of a predetermined period or unit . for example , a school year or a course of study . Periodic assessment measures levels of achievement reached at predetermined intervals throughout the course .Continuous assessment is given to have a continuous updating of judgement about a student's performance .

Approaches to language assessment

The essay-translation Approach: This approach is referred to as the pre-scientific stage of language testing . no special skill in testing is required . there is a lack of concern for objectivity and for test reliability . the test are mainly written exercises : translation , composition ,and grammatical analysis and sometime , an aural / oral component.

frequency in relevant literature.

The structuralist Approach : This approach is mainly characterized by the view that testing can be objective , precise , reliable and scientific .

The Integrative Approach: This approach involves the testing of language in context and is concerned with meaning and the total communicative effect of discourse . Integrative tests are best characterized by the use of cloze testing , which reflects the best feature of the two approaches : analytical and integrative .

The Communicative Approach:

The approach is sometimes linked to the integrative approach as they both emphasize the importance of the meaning of utterance . However , the communicative tests are concerned primarily with assessing real communication . the tests should indicate how well a student can function in the target language .-

Characteristics of Assessment:

Validity: It is the degree to which a test measures what it is supposed a measure, Two questions must always be considered: What precisely does the test measure? How well does it do it?

The types of validity are:

Content validity: it is the extent to which a test adequately the covers the syllabus to be tested. A valid test must be based on a careful analysis of the subject or skill we are testing. **Empirical validity**: the test scores should be related to some independent, outside criteria such as scores given at the end of the course or an external examination. **Face validity**: it is the way the test looks to the examinees, supervisors or the people concerned with the education of the students. **Construct validity**: it is achieved when a test is designed in accordance with a theory of language behavior in learning.

Reliability: it means the stability of test scores. Presumably, if the same test were given twice to the same groups of students, under the same conditions, it would give the same results. The requisites of a dependable test are: Multiple samples; Standard conditions; Standard tasks ; Standard scoring

Practicality: (usability): two parameters are involved:

Economy: the cost in time, money and personnel of administering a particular test.

Ease: the degree of difficulty experienced in the administering and scoring of the test.

Techniques of Assessment

The process of choosing the appropriate technique of testing is governed by a number of considerations including:

a. Subject matter b. Purpose c. age d. ability e. time

Written Assessment

Objective (Closed response)

,*Multiple-choice , True/false, Matching, Rearranging, Labeling , and Supplying.*

Semi-objective

Transformation, completion, close test, short-answer, and questions

Subjective

Essay composition, and précis

A. Objective items: (characteristics)

frequency in relevant literature.

Each item has a precisely defined problem that can be solved in a short time, and has only one predetermined correct answer. An objective test usually has high reliability and content validity.

The test provides wide coverage of subject content. It can be speedily and easily marked and examined statistically. Marking is not influenced by the marker's prejudices.

Multiple-choice items: (the most widely applicable type)

Discrete item: a single complete item. Stem: the initial part of the item in the form of question or incomplete statement. Options: all the choices or answers (usually four or five). Key: the correct answer. Distracters: the incorrect choices.

Advantages of MC items:

The test can have a wide coverage and can be answered in a short time. Large number of learners can be tested in a short time. The test is easy to mark and the marking is objective. It has content validity and high reliability. It is usually free from ambiguity as it restricts the pupils' responses.

Disadvantages of MC items:

Writing a technically sound item is not easy and takes a long time. Pupils' guessing the answers and cheating by copying others' answers can destroy the test results. Only recognition and recall are expected from the test. It can have a negative influence on classwork and learning methods.

It is difficult to find sufficient number of incorrect but plausible distractions.

(8) Suggestions for constructing MC items:

The item should be meaningful and present a definite problem. It should be free from irrelevant material. It should contain only one correct answer. Make sure you do not give away the right answer through grammatical cues. Make the item economical, i.e. words that are repeated in every option. The options should be roughly kept of the same length as one another in a single item. The correct answer should appear in random order. The options should be related to each other (of the same grammatical group). The options have to be arranged clearly. All distracters should be attractive to the student who is not sure of the correct answer.

True / False items (T / F):

They are used for measuring the ability to identify the correctness of statements of facts, definitions of items, statements of principles, etc.....

Advantages of T / F items:

They are frequently easy to construct. They can cover a wide sampling of the course material. Scoring is easy.

Disadvantages of T / F items:

The learning outcomes are largely limited to the knowledge area. The pupils' chances of guessing correctly are high, so that diagnostic value of such a test is weak.

Suggestions for writing T / F items:

Avoid any source of ambiguity, e.g. avoid generalizations and words like: "every, never, nothing, none, often, may, could". Avoid the use of negative statement, especially double negatives. Avoid long complex sentences. Avoid lifting statements verbatim from the textbook.

frequency in relevant literature.

Guessing can be overcome by including a request for an explanation of the selection of a particular judgment.

Matching items:

The test consists of two lists (premises and responses) with instructions as how the matching is to be undertaken. The pupils' task is to identify the pairs of items that are to be associated.

Advantages:

It is possible to measure a large amount of related factual material in a relatively short time. It is constructed easily.

Disadvantages:

They are restricted to the measurement of factual information. Finding significant homogenous material is sometimes difficult. It is also a reducing multiple-choice test.

Suggestions for Constructing Matching Tests:

Use only homogenous material in one single matching test. Let the number of responses be larger than the number of premises. Place all the items for one matching test on the same page. Arrange the responses as far as possible in a logical order. Keep the list of items to be matched brief.

Rearrangement:

It demands the arrangement of a number of words to make a meaningful sentence or to arrange a jumbled series of sentences to formulate a coherent piece of writing. This easy-to-design test can assess effectively the pupils' command of language syntactically and semantically.

Supply Type:

The test is given a group of words followed by a number of incomplete sentences. He is asked to supply the most appropriate words that can fill the blanks.

Labeling :- Is another type of objective tests. The test is required to label certain areas of a diagram or picture.

Semi-Objective Questions:

They combine some of the virtues of both objective and subjective questions. The problems are short and highly structured, but they provide the examinee with the opportunity to compose his own answers. They may require the examinee to complete a sentence or to compose one or more of his according to very specific directions.

Advantages:

The questions are quick, easy to write and less time consuming. They require more than just recognition and recall. They require much less scoring time than would a composition. They demand a certain amount of coherence in the answer. The opportunity for guessing the correct answer is much less than MC.

They take longer to score than objective tests. Quite frequently there are a number of possible right answers. They are not for certain answers (e.g. questions on literary appreciation).

Transformation Items:

They are useful for testing the students' ability to construct and produce sentences. A wide range of grammatical areas can be covered in this test format.

Completion Items:

Carefully constructed, they can test the students' ability to produce correctly grammatical and meaningful sentences. They measure production rather than recognition and easier to construct than MC items.

Close Tests:

frequency in relevant literature.

A text is chosen first, then the tester deletes every n th word. The blanks are of uniform length. The interval at which words are deleted is usually between every fifth and every tenth word. The close test measures the students' ability to decode interrupted messages by making the most acceptable substitutions from all the contextual clues available.

There are two reliable methods of scoring a close test. One mark may be awarded for each acceptable answer or exact answer. Various modified close tests have been proposed such as the multiple choice one or the modified type in which selected words are deleted.

Short-Answer Questions:

They are set in the form of several sub-questions that may be related to the same topic. They could be based on a common set of data that may be in the form of written material, tables, charts, pictures. This technique measures higher mental processes such as reasoning, creative thinking, problem solving and valid conclusions.

C- Subjective (composition) Questions: * Characteristics:

The student must decide on the material which he will include in his answer. Limits are not set upon the length or the style of the response.

Little or no guidance is provided for the test although the question itself should be specific and unambiguous.

Advantages:

The question is relatively easy to set. It is a good measure of power of expression. Composition questions have a desirable influence on study habits. They are often regarded as a device for improving skills.

Disadvantages:

The unreliability of scoring since marking is subjective. The amount of time required for scoring the answer. The limited sampling they provide.

Suggestions for constructing composition questions:

Restrict the use of free response questions to those learning outcomes that cannot be satisfactorily measured by objective items. Clarify the question to make it as specific as possible. Indicate the time limit for each question. Avoid the use of many optional questions (influencing the validity). Use a scoring scheme and prepare an outline of the expected answer. Evaluate all answers to each question at one time.

Oral Assessment

Speaking is a complex requiring the simultaneous use of number of different abilities. Oral assessment, therefore, should cover not only accurate language but also fluency (i.e. ease and speed of the flow of speech) together with comprehension of what is being said.

General – Characteristics of Oral Tests:

They are integrative tests. They are very time-consuming. They are on the whole subjectively scored. There are serious doubts about their reliability. They are considered valid tests.

Suggestions for conducting oral tests:

Warm-up: to make acquaintance of each student and encourage him to perform well by asking him some social questions. **Main task:** to present some tasks to the student and assess his oral. **Roundup:** to tie up an end by encouraging the student to contribute his own individual views on the topics discussed.

Methods of Oral Assessment Tests:

frequency in relevant literature.

Direct speaking tests (oral interview): the examinee is asked to engage in fact-to-fact communicative exchange with the examiner. Indirect tests (paper-and-pencil tests): they do not require any active speech production on the examinee's part. Semi-direct tests: they elicit the examinee's active speech by means of pictures, printed materials or tape recordings.

Techniques for Assessing Oral Skills:

The dialogue

Examiner-student: The examiner selects one of the dialogues in the textbook. He takes one role and the student takes the other role.

Student-student: Two students are given a situation based on one of their textbook dialogues. The teacher then assigns roles..

The interview: It offers a realistic means of assessing the total oral skills in a natural speech situation. It is to be given either in the usual examiner student form (individual interview) which is nowadays strongly recommended because it generates an increased amount of "student talking time" it is useful for the teacher to prepare a series of questions on a wide range of topics, e.g. family, home, school, sports, and holidays etc...

Picture-Based Questions: Pictures provide a realistic context for a brief and structured conversational exchange. The student is given a picture or photograph in a given time. The examiner's version of the picture has a set of suggested questions and follow-up topics. The test is assessed on his command of fluency, grammar, vocabulary, phonology and accuracy of description or narration.

Situations: The student may be told about a certain situation to which he is expected to make suitable response, e.g.

You are at a party and accidentally you spill some juice over someone's white dress / suit. What would you say to her / him?

Reading Aloud: The examiner chooses a suitable passage and gives it to the student to read. The student usually has a few minutes in which to study the text silently before being asked to read it aloud. The examiner should not attempt to test the whole text, but a few specific points, then he has to evaluate the student's ability to read with ease and speed.

Paper-and-Pencil Technique: These tests are to be regarded as an awareness of sound / symbol correspondence rather than tests of pronunciation. Students need to have training to answer such tests during the course before testing them, e.g. Write the number of the item and the letter of the word that carries the given sound:

Criteria for Oral Assessment:

Accuracy: The level of accuracy in grammar, vocabulary and pronunciation.

Appropriate: The use of language broadly appropriate to what the speaker means to say.

Fluency: The ease and speed in speaking

Methodology

This chapter presents comprehensive information about the research methodology employed in the current study. First, the research design will be discussed. After that general information about the participants of the study will be provided. Following that, information about selecting the materials used to collect information will be discussed. Finally, the procedures used to collect data and the procedures for analyzing these data will be highlighted.

frequency in relevant literature.

The current study intends to investigate The importance of using test based paper instead of using copybook test that the minister of Iraqi education test system followed. It also tries to see whether there are any significant differences in the use of these test based papers strategies among students teachers from their attitude . Another aim of the study is to explore the significant of remove the copybook test to help Iraqi government crisis and high budgets of ministry of education in Iraq for education test system for school. In fact, unchecked spending that led to a financial crisis now made worse by low oil prices and high education spending demands.

Research Design

The best method for answering the research questions seemed to be by designing a survey study using a qualitative methods approach. According to Creswell (2003) using qualitative method is “an inquiry strategy that is focused on converging or triangulating different qualitative data sources” (p. 210). Moreover, by using this method it is hoped that more detailed information could be collected and a comprehensive picture of the topic could be observed. Hence, the study involved a qualitative phase (semi-structures interview), semi- structured interviews were employed to discover the awareness of the instructors at this school about test based paper and the strategies that their students used most frequently.

Participants

The group in the current study included five teachers at the Al-Muthan preparatory school for boys . All the participants were male and have B.A degree in English language from Iraqi universities.

- Mr. Ahmed has been teaching English since the early 2000s and he is in his mid- 30s.
- Mr. Ali has been teaching English since 1996 and he is in his late 50s.
- Mr. Salem has been teaching English since late the 1990s and he is in his mid - 40s.
- Mr. Husain has been teaching English since 2000 . He is in his mid- 30s.
- Mr. Saleh has been teaching English since 2006. He is also in his mid- 30s.

To reach these aims, a semi-structured interview schedule was used. Best and Kahn (2006) explain that selection of effective data collection methods are determined on the basis of “what is to be studied” (p. 257).The reasons for choosing the interview as a method to collect data was that this instrument provides more deep and detailed information from an interviewee than what is available through other collection method, such as surveys (Boyce& Neale, 2006). Moreover, it provide comprehensible results from a small number of participants which is not possible by using statistical analysis

Interview is one of main data collection tools used in qualitative studies (Patton, 2002). It is a conversation between the investigator and the participants in his or her study. There are many forms of interviews such as, individual face-to-face interviews, group face-to-face interviews, mailed or self administered questionnaires, and phone surveys (Siriwan, 2007). There are also different styles in interviewing such as fully structured, semi-structured or unstructured interviews (Brown, as cited in Siriwan,

frequency in relevant literature.

2007). The interview used in this study was individual semi-structured face-to-face interview. Patton (1990) defined semi-structured interviewing as “topics and issues to be covered are determined in advance, in outline form; interviewer decides sequences and wording of questions in the course of the interview” (pp. 288-289). This kind of interview provides some freedom for the interviewers to ask and/or give clarification when they feel that the interviewees’ answers are not clear. Also, it helps the interviewer to ask additional questions in order to explore more information about the interviewees’ feelings and opinions and all this information will support his or her main topic (Corbetta, 2003). On the other hand, structured-interviewing means asking the same questions in the same order for all the interviewees (Patton, 1990). According to Patton (1990), these standardized questions don’t give the interviewer flexibility and this could affect the interviewee’s answers. Corbetta (2003) explains that sometimes interviewees don’t understand the questions and this could affect the type as well as the quality of their answers. This had happened in the course of the current study, when the researcher asked two of the participants her first interview question. The instructors didn’t understand the meaning of the questions; therefore, they gave irrelevant answers. Yet, because of following a semi-structured interview format, the researcher was able to clarify the meaning of the questions and further delve into the reasons behind the participants’ initial responses. Thus, the first reason for using semi-structured interview in this study was that fact that such a format provided a very flexible technique, offering useful data when the sample size is small, which was the case in this study. Secondly, it gave the interviewees more freedom and control to express their thoughts (Drever, as cited in Pathake & Intrat, 2012).

Several drafts of the interview questions were written before getting to the final version of the main themes to be discussed. These versions were also checked with the researchers’ supervisor, who is experienced in conducting qualitative research. The key themes that were covered during the interviews contained questions about the strategies the instructors themselves used in test, their educational background about this matter, the strategies that their students used to answer the test based on their observations and expectations, and their opinions regarding any problems that their students may be encountering regarding test (see Appendix C for the interview schedule).

Despite the fact that the participants were encouraged to explain their views in detail by the help of follow up questions such as “Can you tell me a little bit more about that?”, some of them seemed reluctant to provide details and they gave very brief answers to some of the questions. One possible explanation for this could be that they felt that some of the questions in my interview were “knowledge questions.” According to Patton (as cited in Glesne, 2006) knowledge questions can give interviewees a feeling of “being tested”, and in such cases, the interviewees feel “uneasy when they have to say “I don’t know” to a question that you assumed they would know” (p. 82). To avoid giving any such expressions or assumptions to the interviewees that she assumed they should know the answer, the researcher preceded all her knowledge questions with sentences like “Do you think...”, “Based on your observation, what do you think about...?” or “How do you think...”.

All of the interviews were conducted in English and the duration of the interviews varied from five minutes to fifteen minutes. The audio recordings for all the interviews added up to approximately 35 minutes. Each interview was then transcribed using the following transcription conventions (Schegloff, 1987):

Normal –utterances in English

frequency in relevant literature.

() – words or phrase used to make the meaning clear put in parentheses
[laughing] transcriber’s comments are put in square brackets

Analysis

To analyze the data collected through the interviews, open coding suggested by Pattern (2002) was used. In very simple terms, coding means “identify what appears to be important and give it name (code)” (Glesne, 2006, p. 154). According to Pattern (2002), there are three steps in analyzing qualitative data (organizing the data, description, and interpretation). In this study, after reviewing the transcripts of the interviews several times, the data was organized based on the interviewees’ responses for each of the questions in the interview. The interview questions guided the initial analysis. Several themes emerged and these themes were noted down and then grouped based on their frequency. The emergent codes were compared several times with all the remaining interview transcripts to check their validity.

Conclusion

In this chapter, detailed explanations about the research methodology employed in the current study were provided. The rationale for choosing qualitative method approach employed in the current study. and the type of interviews questions used were given. Detailed information about the procedures for collecting and analyzing data were also explained. in the qualitative of the research design were described in detail. The results of the analysis for this study of the study will be presented in the next chapter.

FINDINGS AND DISCUSSION

Presentation

In this chapter, the results from the interviews and observations will be discussed in detail. The results will be presented based on the research questions that guided the current study. Are Iraqi education test system economy to save time and money? How can make the test system more economy? What is the effect of removing the copybook test and replace it by based test paper? According to interviews with both teachers and their students, the results showed that both of them prefer test based paper to save time and easier to answer and correct. All teachers agreed that test based paper more easier that copybook test paper. They believed that test based paper could save time and money both to government and them. The students on the other hand, showed that the test based paper more clear to them and they can follow the question easily and they would never forget or mixed any question.

Moreover, according to my observation to the monthly test exam, the paper easier to save than copybook test exam. The teacher were smoothly correct the paper because they can see the questions directly. The teacher could save more than one hour comparing with traditional of copybook test paper for each test.

According to the ministry of education in Iraq, the ministry spent more than 200,000,000 million copybook test paper for examination. Table (1) shows the spending copybook test paper for each year.

<i>students</i>	<i>Copybook for each one</i>	<i>total</i>	<i>The price for each one</i>	<i>The total amount money each year</i>
-----------------	----------------------------------	--------------	-----------------------------------	---

frequency in relevant literature.

10,000,000	25	250,000,000	250	62,500,000,000
------------	----	-------------	-----	----------------

According to the table above which shows the ministry of education in Iraq spent 62500000000 milers Iraqi dinar for examination each year.

Conclusion

In this chapter, detailed information about the qualitative findings of the study were presented and discussed. The results showed that Al Muthana preparatory school students were happy and they preferred to use test based paper. On the other hand, the qualitative results showed that the teachers at Muthana preparatory school have limited chance to use test based paper because the hard strict policy of the ministry of education in Iraq and supervisors..

The findings suggested that the policy followed by the ministry of education in Iraq about the test and examination should be changed\ a immediately to save time, money, and place to save the copybook test paper. there was a correlation between the strategies that the instructors used in testing and/or advised their students to follow in testing and those that the students preferred/used. The results also showed that learners had an ability to develop their own strategies in testing without any explicit training in this matter, and there were real gap between teaching and testing. In the next and final chapter, general findings of the study will be presented. In addition, implications for pedagogical purposes and recommendation for further research in this area will be provided.

CONCLUSION AND RECOMMENDATIONS

Introduction

This chapter provides a general look at the current study. First, the major findings of the research presented in Chapter Four will be summarized. Then, a set of implications and recommendations will be provided for educational purposes and for possible extension of the present research. At the end, the conclusion of the study will be provided.

Main Findings

This study was designed to identify the important of change the policy of testing to save money and time used by ministry of education in Iraq. The finding showed that ministry of education in Iraq spent more than 62500000000 milers Iraqi dinar for examination each year. However, by using test based paper we can save that amount of money and we can build new schools or provide Iraqi school with computer to develop the electronic test and learning to follow the advanced learning around the world.

Another findings, that the test based paper can save time both to teachers and the students . The students found it easier that the traditional test. The teacher , on the other hand felt happy with this model of testing and they saved at least one hour from their time for each exam correction. Moreover, the results showed that the test based paper

frequency in relevant literature.

was easier to store than the copybook and it could be used the same boxes after written the years and material for each exam to save the paper.

Several key findings emerged from the interviews with the instructors. First, they believed that the Iraqi test system is traditional and old. They understand that the basic things to do is to replace the people in charge to develop new, modern, healthy, and economic test in Iraqi schools.

Conclusion

The Iraqi education system of testing and examination needs to be improved. The modern test should be economic in time and money. The ministry of education in Iraq spent a huge amount of its financial utility on test procedures. The system deviated from the basic principles. Harris, 1969 showed illustrated the good testing

"The profoundly commendable test should be economy in time, effort, and expenses. A good test should not be expensive, very long, and tiring. Ease of Administration and Scoring. There should be full, clear directions so that the test administrator can perform the task quickly and efficiently." (Harris, 1969, p.p. 21)

Recommendations for Further Research

Considering the results and the limitations of this study, several recommendations can be made. First, more detailed information about the Iraqi test system that the ministry of education used, such as the reasons behind choosing specific strategies over others for more than hundred years or the actual time that the teachers employ these strategies in testing, need to be obtained. Therefore, there is a need to triangulate the results presented in the first part of this study (the used of test based paper) with in-depth interviews, think aloud sessions and in-class observations. In other words, there is need for research studies that employ more than a single method (interviews) to identify the level and variety of the test system. Since not all the instructors at the wasit directory of education accepted to be part of this study, the researcher could not explore the awareness levels of all the instructors at this city. Therefore, other studies are needed to cover all the instructors and may be extended to other cities around Iraq in other provinces of Iraq for comparative reasons. In addition, this study depended on conducting only one interview with each instructor to obtain the qualitative data. In order to obtain more information about the awareness levels of the instructors about Iraqi test system, more data using different data collection methods, such as observations of actual teaching sessions or more interviews with each instructor are needed.

Another recommendation that can be made for further research deals with the context and the participants of the study. The current study was conducted with adult learners (preparatory level of the fourth stage). Further research should be conducted with students at other educational levels, such as secondary elementary and primary school students or university level in Iraq. Also, all the participants in this study were drawn from one fourth classes in the south part of Iraq. Therefore further research can be designed to include the remaining parts of Iraq as well.

Since, most of the students who participated of the study were male; the researcher could not carry out any significance tests based on gender. Therefore, studies

frequency in relevant literature.

with more participants of both genders are required to be able to conclude regarding this matter.

This study examined the relationship between test based paper and copybook test paper used by Iraqi education system at Al Muthana preparatory school students and their teachers' about the test system in Iraq. Literature suggests that there is a significant relation between proficiency levels of the Learners and teachers use. In addition no data on the actual proficiency levels of students was collected. Therefore, further research should be conducted to examine the relationship between test paper test and electronic test and other variables, such as proficiency level of the learners. In other words, there is a need to understand the strategies that the ministry of education in Iraq follow in test system students with high and low proficiency level most frequent use based on standardized, reliable, and valid measurement of all participants' proficiency levels.

Conclusion

This study showed that the ministry of education in Iraq, EFL Teachers and the policy makers need to develop their awareness about test and how they develop the test system to be more economic because of Iraqi financial crisis in 2015. The teachers need to look at any subject could help in shaping the learners' knowledge or beliefs about that subject. Moreover, both teachers and students need to update their awareness and knowledge of different learning strategies of testing as the field of language learning is evolving every day and new discoveries, such as internet, are added to help both learners and teachers with their learning/teaching practices. Iraqi education system is old and need to be improve to be more economic and modern by using electronic test.

REFERENCES

- Abazzaz, A.H. (2005). *Investigating the impact of vocabulary instruction on developing reading comprehension*. Unpublished MA thesis. University of Tikrit, Saladin, Iraq.
- Abercrombie, D. (1999) *Elements of General Phonetics Edinburgh: Edinburgh University Press*.
- A Committee in the Ministry of Education (1981). *The New English Course for Iraq Book VII: Teacher's Guide* (1st rev. ed). Baghdad: Ministry of Education.
- Abdul-Rahman, M. H. (1989). "Teachers' Behaviour in Writing Classes: Their Techniques and Responses to College Students' Written Assignments". Unpublished M.A. Thesis: College of Arts, University of Basrah.
- Al-Azzawi, K. K. (1998). "Evaluating EFL Iraqi College Students in Written Composition with Instructional Implications". Unpublished Ph.D. Dissertation: College of Education/Ibn Rushd, University of Baghdad.
- Al-Bahadli, H. J. (2010). "The Impact of Using Dialogue Journals in Developing Writing Skills". Unpublished M.A. Thesis: College of Basic Education, University of Babylon.
- Alderson, J., & Banerjee, J. (2001). "Language Testing and Assessment (Pt.1)". *Language Testing*, 18(4), 213-236.
- Al-Hamash, K.I. (1984). *A Survey of English Textbooks in Primary and Secondary Schools in Iraq*. Baghdad: IDELTI.
- _____, Radhi, A.J., Al-Mutalibi, A.Y., Ahmed, R., Al-Chalabi, S.A., & Al-Khazraji, K.S. (2009). *The New English Course for Iraq Book VII: The Pupil's Book*. Baghdad: Al-Sa'adoun Printing Company.
- Al-Karkhi, A. Y. (1999). "Difficulties Faced by Iraqi EFL College Students in Writing English Composition". Unpublished M.A. Thesis: College of Education/Ibn Rushd, University of Baghdad.
- Al-Khafaji, R. M. (2005). "The Effect of Using the Jig Saw Technique on Writing Composition at University Level". Unpublished M.A Thesis: College of Education for Women, University of Baghdad.
- Al-Temimi, S. H. (2005). "Self-Correction-and-Rewriting Technique: The Instructional Aspect of Composition Test". Unpublished Ph.D. Dissertation: College of Education/Ibn Rushd, University of Baghdad.
- AlZubaidi, H.A (2014) *The effect of using video recording as reflecting teaching tools to improve the performance of EFL teachers in Turkey*. Master thesis .Published at NEU.
- Anastasi, A. (1988). *Psychological Testing*. New York: MacMillan Publishing Company.
- Andrade, H. & Boulay, B. (2003). "Role of Rubric-Referenced Self- Assessment in Learning to Write". *The Journal of Educational Research*, 97(1), 21-34.
- _____, & Du, Y. (2005). "Student Perspectives on Rubric-referenced Assessment". *Practical Assessment, Research & Evaluation*. A peer reviewed electronic journal, 10 (3), 1-11.
- Armstrong, T. (1994). *Multiple Intelligences in the Classroom*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Assessment Reform Group (2002). *Assessment for Learning: 10 Principles*. Cambridge: Cambridge University Press.

- Atkins, M., Beattie, J., & Dockrell, W.B. (1993). *Assessment Issues in Higher Education*. School of Education, University of Newcastle upon Tyne.
- Bachman, L. F. (2004). *Statistical Analyses for Language Assessment*. Cambridge: Cambridge University Press.
- Badger, R., & White, G. (2000). "A Process Genre Approach to Teaching Writing". *ELT Journal*, 34 (2), 153-160.
- Barnett, M. A. (1989). "Writing as a Process". *The French Review*, 63(1), 31-44.
- Best, J. (1981). *Research in Education* (4thed.). New Jersey: Prentice-Hall Inc.
- Black, P. (1998). *Testing: Friend or Foe? Theory and Practice of Assessment and Testing*. London: Falmer Press.
- _____ & William, D. (1998a). "Assessment and Classroom Learning". *Assessment in Education*, 5 (1), 7-74.
- _____. (1998b). "Inside the Black Box: Raising Standards through Classroom Assessment". *Phi Delta Kappan*, 80(2), 139-148.
- Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2003) *Assessment for Learning: Putting it into Practice*. New York: Open University Press.
- _____. (2004). "Working Inside the Black Box: Assessment for Learning in the Classroom". *Phi Delta Kappan*, 9-21.
- Blanche, P. (1988). "Self-assessment of Foreign Language Skills: Implications for Teachers and Researchers". *RELC Journal* 19(1), 75-93.
- Bloom, B., Madaus, G. F., & Hastings, J. T. (1981). *Evaluation to Improve Learning*. New York: McGraw- Hill.
- Blue, G. M. (1988). "Self-assessment: The Limits of Learner Independence". In A. Brookes & P. Grundy (Eds.), *Individualization and Autonomy in Language Learning* (100-118). London: Modern English Publications in Association with the British Council.
- Boud, D. (Ed.) (1995). *Enhancing Learning through Self-Assessment*. London and Philadelphia: Kogan Page.
- Broadfoot, P., Dockrell, B., Gipps, C., Harlen, W. & Nuttl, D. (1992). *Policy Issues in National Assessment: BERA Dialogues n.7*. Clevedon: Multilingual Matters Ltd.
- Brown, F. G. (1976). *Principles of Educational and Psychological Testing* (2nd ed.). New York: Holt, Rinehart & Winston.
- Brown, H. D. (2004). *Language Assessment: Principles and Classroom Practices*. New York: Pearson Education, Inc.
- Brown, J. D. (2005). *Testing in Language Program: A Comprehensive Guide to English Language Assessment*. New York: McGraw-Hill.
- Brown, S. (1990). "Assessment: A Changing Practice". In T. Horton (Ed.), *Assessment Debates*. London: Hodder & Stoughton in Association with the Open University Press.
- _____, Rust, C., & Gibbs, G. (1994). *Strategies for Diversifying Assessment in Higher Education*. Oxford: Oxford Centre for Staff Development.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton Mifflin Company.
- Chamot, A.U., & O'Malley, J.M. (1994). *The CALLA Handbook: Implementing the Cognitive Language Learning Approach*. Reading, MA: Addison Wesley.
- Cho, K., Shunn, C., & Wilson, R. (2005). "Validity and Reliability of Scaffolded Peer Assessment of Writing from Instructor and Student Perspectives". *Journal of Educational Psychology*, 98(4), 891-901.

- Christensen, L. B. (1980). *Experimental Methodology* (2nd ed.). Boston: Allyn & Bacon, Inc.
- Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Cox, C. (2002). *Teaching Language Arts*. Massachusetts: Allyn and Bacon.
- Coyle, D. (2000). "Changing the Rules of the Game: Adolescent Voices Taking Control: If Only They Would, If Only They Could". In B. Sinclair, I. McGrath, & T. Lamb (Eds.), *Learner Autonomy, Teacher Autonomy: Future Directions* (24-36). London: Longman.
- Crooks, T. J. (1988). "The Impact of Classroom Evaluation Practices on Students". *Review of Educational Research*, 58(4).
- Cushing Weigle, S. (2002). *Assessing Writing*. Cambridge: Cambridge University Press.
- Dam, L. (1995). *From Theory to Classroom Practice*. Dublin: Authentik Learning Resources.
- Davies, A., Brown, A., Elder, C., Lumley, T., & McNamara, T. (1999). *Dictionary of Language Testing*. Cambridge: Cambridge University Press.
- Deale, R.N. (1975). *Assessment and Testing in the Secondary School*. London: Evans Bros.
- Dickens, C. (2009). *Oliver Twist: Literary Reader II*. Simplified Version by a Committee in the Ministry of Education. Baghdad: Zarqaa Al-Yamamah Press.
- Dickinson, L. (1987). *Self-Instruction in Language Learning*. Cambridge: Cambridge University Press.
- Divaharan, S., & Atputhasamy, L. (2002). "An Attempt to Enhance the Quality of Cooperative Learning through Peer Assessment". *Journal of Educational Enquiry*, 7(2), 72-83.
- El-Koumy, A. S. (2001). "Effect of Self-Assessment of Writing Processes versus Products on EFL Students' Writing". Egypt: Suez Canal University.
- Erwin, T. D. (1991). *Assessing Student Learning and Development: A Guide to the Principles, Goals, and Methods of Determining College Outcomes*. New Jersey: Jossey-Bass.
- Eschholz, P.A. (1980). "The Prose Models Approach: Using Products in the Process". In T.R. Donovan & B.W. McClelland (Eds.), *Eight Approaches to Teaching Composition*. Urban: National Council of Teachers of English.
- Freire, P. (1970). *Pedagogy of the Oppressed*. New York: Herder & Herder.
- Gardner, D. (2000). "Self-Assessment for Autonomous Language Learners". *Links and Letters*, 7, 49-60.
- Gardner, J. (2006). *Assessment and Learning*. London: Sage.
- Gibbs, G. (2003). "Using Assessment Strategically to Change the Way Students Learn". *Assessment Matters in Higher Education: Choosing and Using Diverse Approaches*. Buckingham: SRHE and Open University Press.
- _____ & Habeshaw, T. (1990). "An Introduction to Assessment: Induction Pack 4". *Birmingham, Standing Conference on Educational Development* (SCED Paper 57).
- Gipps, C. (1994). *Beyond Testing: Towards a Theory of Educational Assessment*. London: The Falmer Press.

- Glass, G. V., & Stanley, J. (1970). *Statistical Methods in Education and Psychology*. New Jersey: Prentice-Hill.
- Goal Based Education Program (1984). *Effective Schooling Practices: A Research Synthesis*. Portland: Northwest Regional Educational Laboratory.
- Graves, D. H. (1978). *Balance the Basics: Let Them Write*. New York: Ford Foundation Papers on Research about Learning.
- Hairston, M. (1982). "The Winds of Change: Thomas Kuhn and the Revolution in the Teaching of Writing". *College Composition and Communication*, 33(1), 76 – 88.
- Halliday, M. (1978). *Language as a Social Semiotic: The Social Interpretation of Language and Meaning*. London: Edward Arnold.
- Halsted, I. (1975). "Putting Error in its Place". *Journal of Basic Writing*, 1 (1), 72-86.
- Hamayan, E. V. (1995). "Approaches to Alternative Assessment". *Annual Review of Applied Linguistics*, 15, 212-226.
- Harlen, W. (2006). "On the Relationship between Assessment for Formative and Summative Purposes". In J. Gardner (Ed.), *Assessment and Learning* (103-117). London: Sage
- _____ & Crick, R. D. (2003). "Testing and Motivation for Learning". *Assessment in Education: Principles, Policy and Practice*, 10(2), 169 – 207.
- Harris, D. P. (1969). *Testing Language as a Second Language*. New York: McGraw-Hill Book Company.
- Harris, M. (1997). Self-Assessment of Language Learning in Formal Settings. *ELT Journal*, 51(1), 12-20.
- _____ & McCann, P. (1994). *Assessment: Handbooks for the English Classroom*. Oxford: Heinemann.
- Harrocks, J. E., & Schonover, T. I. (1989). *Measurement for Teachers*. New York: Charles E. Merrill Publishing Company.
- Heaton, J.D. (1975). *Writing English Language Tests*. London: Longman.
- Herman, J. J., Aschbacher, P. R., & Wintners, L. (1992). *A Practical Guide to Alternative Assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Heron, J. (1988). "Assessment Revisited". In D. Boud (Ed.), *Developing Student Autonomy in Learning* (77-90). London: Kogan Page.
- Holec, H. (1981). *Autonomy and Foreign Language Learning*. Oxford: Pergamon.
- _____, Little, D., & Richterich, R. (1997). *Strategies in Language Learning and Use*. Strasbourg: Council of Europe.
- Huerta-Macias, A. (1995). "Alternative Assessment: Responses to Commonly Asked Questions". *TESOL Journal*, 5, 8-10.
- Isaac, S., & Michael W. B. (1977). *Handbook in Research and Evaluation*. California: Robert R. Knapp.
- Janssen-van Dieten, A. (1989). "The Development of a Test of Dutch as a Foreign Language: the Validity of Self-Assessment by Inexperienced Subjects". *Language Testing* 6(1), 30-46.
- Jassim, F. Y. (1988). "Problems of Teaching Composition in the New English Course for Iraq at the Secondary Stage: Diagnosis and Remedy". Unpublished M.A. Thesis: College of Arts, University of Basrah.
- Javaherbakhsh, M. R. (2010). "The Impact of Self-Assessment on Iranian EFL Learners' Writing Skill". *English Language Teaching*, 3(2), 213-217.

- Jiliang, C., & Kun, W. (2007). "An Investigation of Scorer Reliability of Self and Peer Assessment of EFL Writing among Chinese College Students". *CELEA Journal*, 30(1), 3-11.
- Klenowski, V. (1995). "Student Self-evaluation Processes in Student-centred Teaching and Learning Contexts of Australia and England". *Assessment in Education*, 2(2), 145-163.
- Kohonen, V. (1992). "Foreign Language Learning as Learner Education: Facilitating Self-direction in Language Learning". In B. North (Ed.), *Transparency and Coherence in Language Learning in Europe* (71-87). Strasbourg: Council of Europe.
- _____. (1997). "Authentic Assessment as an Integration of Language Learning, Teaching, Evaluation and the Teacher's Professional Growth". In A. Hutha, V. Cohonen, L. Kurki-Suonio & S. Louma (Eds.), *Current Developments and Alternatives in Language Assessment* (7-22). Jyvaskyla: University of Jyvaskyla.
- _____. (1999). "Authentic Assessment in Affective Foreign Language Education". In J. Arnold (Ed.), *Affect in Language Learning* (279-294). Cambridge: Cambridge University Press.
- Kurt, G., & Atay, D. (2005). "The Effects of Peer Feedback on the Writing Anxiety of Prospective Turkish Teachers of EFL". *Journal of Theory and Practice in Education*, 3(1), 12-23.
- Lado, R. (1961). *Language Testing: The Construction and Use of Foreign Language Tests*. New York: McGraw-Hill Inc.
- Lantolf, P., & Appel, G. (Eds.) (1994). *Vygotskian Approaches to Second Language Research*. Norwood, N.J.: Ablex Publishing Corporation.
- Lewin, M. (1979). *Understanding Psychological Research*. New York: John Wiley & Sons.
- Linse, C. T. (2005). *Practical English Language Teaching: Young Learners*. New York: McGraw-Hill.
- Little, D. (1991). *Learner Autonomy 1: Definitions, Issues and Problem*. Dublin: Authentik.
- _____. (1997). "Strategic Competence Considered in Relation to Strategic Control of the Language Learning Process". In H. Holec, D. Little, & R. Richterich (Eds.), *Strategies in Language Learning and Use* (9-132). Strasbourg: Council of Europe.
- _____. (1998). "Why Focus on Learning Rather than Teaching". In D. Little, L. Dam, & J. Timmer (Eds.), *Focus on Learning rather than Teaching: Why and How?* (3-17). Dublin: Trinity College, Centre for Language and Communication Studies.
- _____. (2000). "Learner Autonomy and Human Interdependence: Some Theoretical and Practical Consequences of a Social-Interactive View of Cognition, Learning and Language". In B. Sinclair, H. McGrath & T. Lamb (Eds.), *Learner Autonomy, Teacher Autonomy: Future Directions* (15-23). Harlow: Longman.
- _____. (2002, September). "We're All in it Together: Exploring the Interdependence of Teacher and Learner Autonomy". In L. Karlsson, F. Kjisik & J. Nordlund (Eds.) (2002), *All Together Now* (45-56) (Papers from the 7th Nordic Conference and Workshop on Autonomous Language Learning). Helsinki: University of Helsinki, Language Center.

- _____. (2003). "Learner Autonomy and Public Examinations". In D. Little & E. Unshod (Eds.), *Learner Autonomy in the Foreign Language Classroom: Teacher, Learner, Curriculum and Assessment* (223-233). Dublin: Authentik.
- _____. (2005). "The Common European Framework of Reference for Languages and the European Language Portfolio: Involving Learners and their Judgments in the Assessment Process". *Language Testing*, 22, 321-336.
- Lynch, B. K. (2001). "Rethinking Assessment from a Critical Perspective". *Language Testing*, 18(4), 351-372.
- Madsen, H. (1983). *Techniques in Testing*. Oxford: Oxford University Press.
- Mahon, T. (1992). "From Sentence to Story: A Process Approach to the Development of Composition Skills in the Primary School". In M. Lau & M. Murphy. (Eds.), *Developing Writing: Purposes and Practices*. Hong Kong: Institute of Language in Education.
- Malehorn, H. (1994). "Ten Measures Better than Grading". *Clearing House*, 67(6).
- Mcdevitt, D. (1989). "How to Cope with Spaghetti Writing". *ELT Journal*, 43(1), 19-23.
- McNamara, T. (2000). *Language Testing*. Oxford: Oxford University Press.
- Mercer, N. (2000). *Words and Minds: How We Use Language to Think Together*. London: Routledge.
- Mok, M., Lung, C., Cheng, D., Cheung, R., & Ng, M. (2006). "Self-Assessment in Higher Education: Experience in Using a Metacognitive Approach in Five Case Studies". *Assessment & Evaluation in Higher Education*, 31(4), 415-433.
- Morocco, G., & Soven, M. (1990). "Writing across the Curriculum in the Foreign Language Class: Developing a New Pedagogy". *Hispania*, 73, 845-849.
- Mousavi, S. A. (1999). *A Dictionary of Language Testing* (2nd ed.). Tehran: Rahnama Publications.
- Myhill, D. (2005). "Testing Times: the Impact of Prior Knowledge on Written Genres Produced in Examination Settings". *Assessment in Education* 12(3), 289-300.
- Ngar-Fun, L., & Carless, D. (2006). "Peer Feedback: the Learning Element of Peer Assessment". *Teaching in Higher Education*, 11(3), 279-290.
- Nunan, D. (1988). *The Learner-centred Curriculum*. Cambridge: Cambridge University Press.
- _____. (1996). "Towards Autonomous Learning: Some Theoretical, Empirical and Practical Issues". In R. Pemberton, S.L. Edward & H.D. Pierson (Eds.), *Taking Control: Autonomy in Language Learning* (13-26). Hong Kong: Hong Kong University Press.
- _____. (2001). *Second Language Teaching and Learning*. Boston: Heinle and Heinle Publishers.
- _____. (2003). *Practical English Language Teaching*. New York: McGraw-Hill.
- O'Malley, J.M., & Valdez Pierce, L. (1996). *Authentic Assessment for English Learners*. New York: Addison-Wesley Publishing Company.
- Oller, J. W. (1979). *Language Test at School*. London: Longman.
- Olsher, D. (1996). *Words in Motion: An Interactive Approach to Writing*. Oxford: Oxford University Press,

- Orsmond, P., Merry, S., & Reiling, K. (2000). "The Use of Student-derived Marking Criteria in Peer and Self-Assessment". *Assessment & Evaluation in Higher Education*, 25 (1), 23-38.
- Oscarson, M. (1978). *Approaches to Self-assessment in Foreign Language Learning*. Council of Europe: Pergamon Press.
- _____. (1989). "Self-Assessment of Language Proficiency: Rationale and Applications". *Language Testing*, 6, 1-13.
- _____. (1997). "Self-Assessment of Foreign and Second Language Proficiency". In C. Clapham & D. Corson (Eds.), *Encyclopedia of Language and Education: Language Testing and Assessment* (7,175-187). Dordrecht: Kluwer Academic Publishers.
- Parson, G. (1985, June). *Hand in Hand: The Writing Process and the Microcomputer. Two Revolutions in the Teaching of Writing*. A Manual for Secondary Teachers. Juneau, AK: Alaska State Department of Education.
- Pennington, M.C. (1995). "The Teacher Change Cycle". *TESOL Quarterly*, 29(4), 705-731
- Peterson, S.S. (2010). "Improving Student Writing: Using Feedback as a Teaching Tool. *What Works? A Research-into-Practice Series Produced by a Partnership between the Literacy and Numeracy Secretariat and the Ontario Association of Deans of Education/University of Toronto*, 29, 1-4.
- Pierce, B.N., Swain, M., & Hart, D. (1993). "Self-Assessment in two French Immersion Programs". *Applied Linguistics*, 14(1), 25-42.
- Pincas, A. (1962). "Structural Linguistics and Systematic Composition Teaching to Students of English as a Second Language". *Language Learning*, 12, 185 – 194.
- Race, P., Brown, S., & Smith, B. (Eds.) (2005). *500 Tips on Assessment* (2nd ed.). London: Routledge Falmer.
- Raimes, A. (1983). *Techniques in Teaching Writing*. New York: Oxford University Press.
- Rea-Dickens, P. (2000). "Classroom Assessment". In T. Hedge (Ed.), *Teaching and Learning in the Language Classroom*. Oxford: Oxford University Press.
- _____ & Rixon, S. (1997). "The Assessment of Young Learners of English as a Foreign Language". In C. Clapham & D. Carson (Eds.), *Encyclopedia of Language and Education: Language Testing and Assessment* (7). Dordrecht: Kluwer Academic Publishers.
- Richards, J. C., & Renandya, W. A. (Eds.) (2002). *Methodology in Language Teaching: An Anthology of Current Practice*. Cambridge: Cambridge University Press.
- _____ & Rodgers, S. T. (2001). *Approaches and Methods in Language Teaching*. Cambridge: Cambridge University Press
- _____ & Schmidt, R. (2002). *Dictionary of Language Teaching and Applied Linguistics*. London: Pearson Education Ltd.
- Rolheiser, C., & Ross, J. (2000). "Student Evaluation - What Do We Know?" *Orbit*, 30 (4), 33-36.
- Sadler, D. R. (1989). Formative Assessment and the Design of Instructional Systems. *Instructional Science*, 18, 119-144.

- Sadler, P. M., & Good, E. (2006). "The Impact of Self- and Peer-Grading on Student Learning". *Educational Assessment* 11(1), 1-31.
- Şahin, S. (2007). "An Application of Peer Assessment in Higher Education". *TOJET*, 7(2), 1, 5-10.
- Salomon, G. (1993). "No Distribution without Individuals' Cognition: a Dynamic Interactional View". In G. Salomon (Ed.), *Distributed Cognitions: Psychological and Educational Considerations* (111-138). Cambridge: Cambridge University Press.
- Sealey, J. (1987). "Peer Editing Groups". *R & D Interpretation Service Bulletin* 9(2).
- Shuangxiang, Z. (2008). "Assessing the Impact of Peer Revision on English Writing of Tertiary EFL Learners". *CELEA Journal*, 31(2), 47-54.
- Silva, T. (1990). "Second Language Composition Instruction: Developments, Issues, and Directions in ESL". In B. Kroll (Ed.), *Second Language Writing: Research Insights for the Classroom* (11-23). Cambridge: Cambridge University Press.
- Smith, D. L., & Hatton, N. (1993). "Reflection in Teacher Education: a Study in Progress". *Education Research and Perspectives* 20(1), 13-23.
- Smith, K. (1999). "Language Testing: Alternative Methods". In B. Spolsky (Ed.), *Concise Encyclopedia of Educational Linguistics* (703-706). Amsterdam: Elsevier.
- Spolsky, B. (1992). "Diagnostic Testing Revisited". In E. Shohamy & R.A. Walton, R.A., (Eds.), *Language Assessment and Feedback: Testing and Other Strategies* (29-39). National Foreign Language Center. Dubuque, IA: Kendall/Hunt Publishing Co.
- Stefani, L. (1998). "Assessment in Partnership with Learners". *Assessment & Evaluation in Higher Education*, 23(4), 339-350.
- Taras, M. (2002). "Using Assessment for Learning and Learning from Assessment". *Assessment & Evaluation in Higher Education*, 27(6), 501-510.
- Taylor, B. P. (1981). "Content and Written Form: A Two-way Street". *TESOL Quarterly*, 15, 5-13.
- Taylor, M. E. (2003). "Using Collateral Material to Improve Writing Performance". *ELT Journal*, 57(2), 149-157.
- Thorndike, R. M., & Dinnel, D.L. (2001). *Basic Statistics for the Behavioral Sciences*. New York: Prentice-Hall Inc.
- Topping, K. (1998). "Peer-assessment between Students in Colleges and Universities". *Review of Educational Research* 68(3), 249-276.
- Torrance, H. (2007). "Assessment as Learning? How the Use of Explicit Learning Objectives, Assessment Criteria and Feedback in Post-Secondary Education and Training Can Come to Dominate Learning". *Assessment in Education: Principles, Policy & Practice*, 14(3), 281 – 294.
- Tudor, I. (1996). *Learner-Centeredness as Language Education*. Cambridge: Cambridge University Press.
- Valdez Pierce, L., & O'Malley, J.M. (1992). *Performance and Portfolio Assessment for Language Minority Students*. Washington, DC: National Clearinghouse for Bilingual Education.
- Vygotsky, L. (1978). *Mind in Society: The Development of Higher Psychological Processes*. London: Harvard University Press.

- Weeden, P., Winter, J., & Broadfoot, P. (Eds.) (2002). *Assessment*. New York: Routledge Falmer.
- Wei, Y., & Chen, Y. (2003). "Supporting Chinese Learners of English to Implement Self-Assessment in L2 Writing". *Proceedings of the Independent Learning Conference 2003, Published on 20th September 2004*.
- Zamel, V. (1982). "Writing: The Process of Discovering Meaning". *TESOL Quarterly*, 16(2), 195- 209.
- _____. (1983). "The Composing Processes of Advanced ESL Students: Six Case Studies". *TESOL Quarterly*, 17, 165- 187.
- Zemach, D., & Islam, C. (2005). *Paragraph Writing: From Sentence to Paragraph*. Oxford: Macmillan Education. .

إلغاء الدفتر الامتحاني للصفوف الغير منتهية

الباحث

م.م. حيدر علي عبد الرضا الزبيدي

إعدادية المثنى للبنين-تربية واسط

ماجستير مناهج واختبارات وطرق تدريس اللغة الانكليزية-تركيا

07801934616

alsadeercom@yahoo.com

بعد إجراء بحث نوعي من قبل الباحث المختص بمناهج واختبارات وطرق تدريس اللغة الانكليزية في إعدادية المثنى ولمدة عام كامل اثبت بأهمية إلغاء الدفتر الامتحاني من الصفوف الغير منتهية واستبداله بالإجابة بنفس أسئلة الامتحان وقد تين من البحث الدقيق أهمية هذه العملية بالنقاط الآتية:

1- الجدوى الاقتصادية توفير مبالغ هائلة لوزارة التربية وذلك بإلغاء الدفتر الامتحاني وتوفير هذه الأموال لدعم برنامج الامتحان الالكتروني. وقد تبين بان الطالب العراقي يحتاج إلى مالا يقل عن 25 دفتر امتحاني سنويا وإذا فرضنا بان عدد طلاب العراق

للفصول الغير منتهية هو 10000000 مليون طالب فهذا يعني توفير 250000000 مليون دفتر سنويا يعني توفير 6250000000 مليار دينار عراقي. جدول رقم واحد يبين الجدوى الاقتصادية من البحث

	عدد الطلبة في العراق	10000000 مليون طالب
احتياج الطالب السنوي للدفتري الامتحاني	25 دفتري	نصف السنة 10-نهائي السنة 10- الدور الثاني 5
عدد الدفتري المستهلكة في العراق سنويا	250000000 مليون	
سعر الدفتري مع النقل والتلف وسوء الخزن	250 دينار عراقي	
المبلغ الكلي	62500000000 مليار دينار عراقي سنويا	

- 2- سهولة التصليح : اتضح من الدراسة بان التصحيح في نفس الورقة يوفر جهد ووقت يصل إلى أكثر من 40% من الوقت والجهد المقدم من قبل المدرس او اللجان التصحيحية
- 3- تقليل الأخطاء في التصحيح: اتضح من خلال البحث المقدم بان السهو من قبل الطالب في تقديم وتأخير الأسئلة وعدم الترتيب يؤثر سلبيًا على درجة الطالب ويؤدي إلى ارتباك المصحح وضياع الوقت بين النظر إلى الأسئلة والدفتري الامتحاني . بينما التصحيح المباشر للورقة الامتحاني يؤدي إلى عدم الوقوع بالأخطاء من قبل الطالب من ناحية والمدرس من ناحية أخرى يعني التصحيح يكون أكثر دقة وعدلا.
- 4- توفير أماكن خزن الدفاتري الامتحانية: إن تراكم الدفاتري الامتحانية وصعوبة ترتيبها وخبزها يؤدي إلى ارتباك في الوصول إليها وتأخذ حيزًا كبيرًا من المساحة في المدارس علما إن مدارس العراق أساسًا تواجه مشاكل بنقص الصفوف نتيجة الأعداد المتزايدة من الطلبة وقلة المدارس. بنما تسهل عملية خزن ورقة الأسئلة لأنها لا تأخذ حيزًا كبيرًا ويمكن الاستفادة من كراتين الورق نفسها لوضع الأوراق المصححة وتكتب عليها اسم المادة والسنة والدور مما يوفر بيئة نظيفة غير استهلاكية ومعالجة النفايات.
- 5- تطوير نمط الامتحان العراقي: قد تقود هذه العملية إلى تغيير نمط الامتحان العراقي مرتبطًا بالعالم لتكون هذه الخطوة المهدي للامتحان الالكتروني الذي يعمل عليه الباحث حاليًا مما يوفر مبالغ الدفاتري والأوراق واللجان الامتحانية ويكون أكثر دقة ومصداقية بالتصحيح الالكتروني الذي يجنب الطالب الاعتراض عليه مونه يصحح الكترونيا.
- 6- تطوير قدرات المدرس والمعلم العراقي لفهم مبادئ التقويم والاختبار حيث ان كل النظريات تقول يجب ان يكون الامتحان اقتصادي بالوقت والأموال والجهد لذلك هذه

العملية سوف تعلم المدرس والمعلم كيفية الاختبار اقتصاديا ونوعيا وتجنب الإسهاب في الامتحانات.

7- ثبت إن الورقة الامتحانية تقلل الجهد النفسي للطالب على عكس الدفتر الامتحاني الذي يربك الطالب مما يؤدي إلى نسيان المواد التي قرأها. وثبت البحث إن الطالب يتمتع أكثر بالورقة الامتحانية وكانت نتائج الطلبة أكثر بايجابية من الطلبة الممتحنين بالدفتر الامتحاني.

التوصيات والمقترحات

- 1- ممكن تطبيق الورقة الامتحانية لتشمل الصفوف المنتهية أيضا
- 2- يجب ان تشمل جميع المواد الدراسية
- 3- تطبق في الجامعات العراقية والتعليم العالي ومؤسسة المعاهد الفنية
- 4- تطوير وتدريب كوادر مختصة في الامتحانات ويفضل ان يكون من حملة الشهادات العليا باختصاص مناهج وطرق تدريس وخريج بلد اجنبي ليحمل أفكار البلد بالتقويم والامتحانات
- 5- أصبح ملح جدا نتيجة الظروف الاقتصادية في العراق بتطبيق الامتحان الالكتروني لتوفير مليارات الدنانير من طبع الأسئلة واللجان والأجهزة والدفتر الامتحاني علما إن لدى الباحث نظام الامتحان الالكتروني (أوفلاين) والذي ممكن تطبيقه بالعراق ويساعد على إن يكون الامتحان شرعي ويجنب مشاكل ظهور الأسئلة الامتحانية وعدالة النتائج كون التصحيح الالكتروني.
- 6- تطبيق المعايير العلمية في اختيار القيادات التعليمية ويجب إن يكون مختص بطرق التدريس او خريج قيادة تربوية مختصة لتطوير التعليم والامتحان بالعراق .
- 7- الاعتناء بالكفاءات العراقية المختصة ويجب إشراكها في عملية بناء التعليم بالعراق حيث لدينا قابليات حقيقية مهمة من قبل وزارة التربية ومنهم من ساهم في تطوير التعليم في بلدان اجنبية متطورة.

الباحث

م.م. حيدر علي عبد الرضا الزبيدي

إعدادية المثنى للبنين-تربية واسط

ماجستير مناهج واختبارات وطرق تدريس اللغة الانكليزية-تركيا

07801934616

07733356449

alsadeercom@yahoo.com