People's Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

Mohamed Kheider University of Biskra

Faculty of Letters and Languages

Department of Foreign Languages

Branch of English

Coming To Grips with The Current Testing Methods:

Strategies To Construct Useful Tests

An Evaluative study of the Didactics Test sheets of First Year Master EFL Students

Dissertation Submitted to the Department of Foreign Languages as Partial

Fulfilment for the Master's Degree in Sciences of Language

Presented by: Boussam Chaouki   **Supervised by**: Mrs. Benidir Samira

Board of Examiners

President: Dr. Saihi Hanan   University of Biskra
Supervisor: Mrs. Benidir Samira   University of Biskra
Examiner: Mrs. Salhi Ahlem   University of Biskra

June 2015

DEDICATION

This work is dedicated to

The ones that gladly brought me into life: My dear Mother and Father

My beloved Sisters and Brothers

My precious family: Antes and Uncles

The ones that I've been blessed to know

My friends group six (2010)

Those who sincerely supported me with their prayers, kindness, and efforts

To all of you, I gladly dedicate this work and these words

Thank you

Acknowledgment

This work would not have been completed without God's help,guidance, and enlightens, for God I owe the greatest thanks.

I am tremendously thankful to my supervisor Mrs. Binidir for her sustained efforts until the end of this work, her patience, and her resourceful advice.

I am also indebted to Mrs. Salhi and DR. Saihi for their time, patience, and suggestions for the betterment of this work.

I am profoundly thankful to Douida Brahim, Bechar Mohamed Amine and Laadjal Karima for helping me finish this research.

I am indefinitely thankful for those teachers who were sincerely caring , giving , and understanding throughout my whole educational life .

I will be always thankful to all of you,

and to all of you I say

THANK YOU

Abstract

This research was undertaken in order to determine whether the TEFL teachers of English branch of Biskra University respect the rules of designing tests or they only rely on their experience. The research conducted in 2014-2015 aim to analyze a sample of tests that were made by didactic teachers in different academic years. We also gathered data about the students' attitude toward the didactic tests through directing a questionnaire to sample of master one student. In order to cover all aspects in our research we also directed an interview to didactic teachers to have a better understanding about the way they design their tests. The document analysis showed that some of the teachers do take in consideration the rules of designing tests however, the interview that was directed to those teachers shows that they rely on their experience more than rules made to suit other student who have different facilities and different circumstances. Even though the students' questionnaire can partially prove that some of these tests are very useful and suitable in our department, it also shows that these tests need to be re-evaluated and designed to meet with the current circumstances. This research was taken to cover only two types of tests, this was due to all the tests in the last 3 years were the same which was an issue in our study process.

Table of content

List of tables

List of Abbreviations and Symbols

TEFL: Teaching English as a foreign language

EFL: English as a Foreign Language

LTAL: Language Teaching and Applied Linguistics

W,Y,X and Z: Variables

I.   Introduction

.   Learning foreign language is a matter of not only studying rules, but also applying them in different context in order to use them easily because the more we practice the better we get feedback and the better we learn. To apply these rules, we need different tasks and activities which differ according to the skill that we would like to improve. Thus, teachers have to provide learners with different tasks to measure what they have gained before. Tests are the most used and the most effective materials that can measure the students' abilities, skills and knowledge. In our work, we tend to identify and differentiate approaches to language testing, to explain their strengths and weaknesses and to discuss the various language testing types, purposes and techniques.

II.   Statement of the problem

The majority of students consider that testing is a vehicle that allows them to pass from one grade to another while researchers view testing as a means which helps us to promote our knowledge. This paper will analyses the types of testing adapted by teachers in the Branch of English studies at the Department of Foreign languages of Biskra, especially in didactics, and the relationship with teaching-learning process. It is observable to the researchers in the field of testing that the tests which are given to the students in the Branch of English at the Department of foreign Languages of Biskra through analyzing some samples of exams of deferent grades (third year and first year master student in our case) that tests focus on the inputs that was discussed in the classroom, however, in the case of didactics (the art of teaching) teachers are neglecting the most important feature of this module which is solving problems that may encounter the teaching-learning process in the classroom. The problem is the choice of an appropriate type of exercises in the test that

are relevant to didactics interest, the students' needs and the objective of the didactics teaching process.

III.   Significance of the study

Testing in university is very important to keep students on trial or to evaluate the materials that are being taught; furthermore, it is to some extent another lesson in the form of case of study to improve the student knowledge. The significance of this study appears in evaluating the testing methods that are used and how the students react to them. In our case, we will be focusing on the didactics module, how tests in this module are constructed and how student react toward them.

IV.   Aims/objectives

This work aims to:

1.   Evaluate the testing methods that are used in didactics.

2.   Investigate whether the basic principles and rules of testing are respected or not.

3.   Identify the main advantages and disadvantages of the tests in didactics module.

V.   Research questions

This dissertation relies on the following questions:

1.   Do the contents of the test serve the aims of didactics?

2.   What are the disadvantages of the testing methods that are used in didactics?

3.   How can the appropriate contents in a test improve the student's awareness to the real problems of teaching-learning process?

VI.   Hypotheses

1.   We hypothesize that the exercises used in the didactic tests are only focusing on recalling what has been discussed in the classroom.

2.   We assume that didactic tests do not obey the basic principle of testing.

VII.   Methodology

We attend to provide hints about our research methodology, sample and tools.

VII.1. The choice of the method

        We intend to use the qualitative method as a design to gather data for this thesis. Furthermore, we plan to derive information from any relevant material to testing field of interest which is a new and a fresh area in the Branch of English at Department of foreign Languages of Biskra as this thesis is concerned. We will take two samples of didactic tests to work on and analyze them according to the rules that guide the teacher while designing a test, in addition, we will hand a questionnaire to master one English students to gather information about their attitude toward the didactic tests, we will also interview some of didactic teacher for the sake of gather information about how they design tests and how they evaluate them. The results of the questionnaire, the interview and the document analysis will be included in the thesis.

VII.2. Population/sample

        We will choose at random group master one English students in the university. This choice is for the reason that these students have been exposed to the methods that are used by the teachers of didactics and probably they have faced the phenomenon that we are studying.

VII.3. Research tools

In this research we attend to take three tools of research. First, the document analysis in order to compare the tests that are constructed by the didactic teachers, with the rules and principles that was made by testing researchers. Second, we attend to direct a questionnaire to the students to gather data about their opinion about the tests. Finally, we will interview the didactic teachers and ask how they design their tests.

VIII.   Research limitations

The research will focus on:

1.   The findings to be reached in this study will be limited in the time and population.

2.   Finding out effective strategies to improve the teaching skills through the tests.

3.   This study will focus on the evaluation of didactic tests

Chapter one: The concept of Testing

Introduction

Testing is no longer only grading the student to pass from one level to the next, but it becomes very important to evaluate the curriculum itself in addition to identifying the weaknesses of the students and determining what they did not master yet. For this sake, it is crucial that tests should deal with all skills (listening/speaking, reading /writing, grammar/vocabulary and so on).

Language testing is relatively a new field in the context of applied linguistics. This field rooted to applied linguistics language testing primarily deals with English Language learners (test-takers) and test designers (teachers, researchers and publishers).The field of language testing therefore consists of four major component: the first one focuses on the "What" also known as the "trail"; these concepts refer to the materials need to be assessed. The second element in these components which is basically relevant to the teacher is the "how" also known as the "method", the teacher or the instructor takes some procedures and strategies to insure constructing good test. The third one is the "test-taker", in our case "the learner", and his performance in the English language test. The last one is the shift from focusing on the test to focusing on why students did/did not master a given material. (Fulcher and Davidson, 2007)

Traditionally, the four components motioned before differ according to the skill being tested. The "What "and the "How" are the most influencing component; they are strongly related to the skills and also have great effect on the other Language testing component; the instructor needs to specify the needed "what" and create the best way or method "how" to test it. Basically, when the designer or the teacher attempts to formulate a test he/she takes in consideration whether the test will focus on only one specific skill (e.g.

oral expression we focus on only listening/speaking) or the test will put all the skills under evaluation (e.g. the secondary school tests, the final mark based on testing all the skills). On the other hand, the "What and How" is influenced by the "test-taker"; the test designer must put in mind whether a given language (in our case English Language) is being taught as a first or a second language.

To sum up, the field of language testing is built on theories and definitions from linguistics, applied linguistics, language acquisition and language teaching as well. Researcher, test designer and teachers have investigated and developed so far good methods that help to enhance language learning. In this chapter we attempt to discuss the differences between the three terms (assessment, evaluation and testing) in order to make a clear distinction for the sake of building a strong basis for our study.

1.1. Key concepts to language testing

When dealing with this subject, one of the core points of confusion that researchers (linguists, teachers and/or learner) encounter is related to terminology of educational assessment. They will find three terms (evaluation, assessment and testing). In some sources, these terms, are used interchangeably to refer to the process f measuring students' skills, ability and knowledge .While in other references, each term has its own definition and use. In this part of our work, we are trying to make the distinction or to show the difference between these terms.

1.1.1. Determine the concept of Evaluation

The word evaluation can be viewed as a general term, we will try to define it and determent it main use in learning process.

1.1.1.1. Definition Evaluation

Evaluation, as Dikins and Germaine stated in their book titled "Evaluation", is an intrinsic part of teaching and learning. It is important for teachers because it provides them with useful information that can help them for future direction of classroom practice, for the planning of courses, and for the management of learning tasks and students (1993, p.17). According to Douglas (2000) and Lynch (2003), the term "Evaluation" refers to the judgment made about a test taker's ability, knowledge or intelligence based on his or her test performance (cited in Deygers, p.10). However, these definitions have limited the evaluation in the educational field which is only a small part of the general meaning of evaluation. In fact, evaluation is a natural activity which we use to judge our deeds in our daily life or the comments that we make when we are asked to take chances. Any judgmental activity we make in our daily life is also known as evaluation but researchers refer to it as "informal evaluation".

1.1.1.2. Evaluation in use

In general, evaluation can be divided into two types: "formal" and "informal". Basically, in the academic interest, the term evaluation will be related to the formal educational purposes only. Evaluation can be divided into four main types as Dikins and Germaine (1993) claimed in their book "Evaluation":

1.1.1.2.1.   Evaluation in education

Evaluating different aspects of the teaching-learning process makes it important to explicit the criteria used in our judgment. The evaluation must be reliable, valid, fair and informative. For this sake, it is crucial to use suitable methods to make decision about teacher work or the learning tasks used in the classroom. Furthermore, the evaluation used

to measure the teacher's performance in classroom differs from the one used to measure learner's performance during teaching –learning process.

1.1.1.2.2.   Evaluation and making decision

Making judgment in our daily life surely differs from the one we make in educational context. The evaluation in education does not affect only an individual but the whole process of education. Making decision about a student pass a grade or a teacher promotion has a higher perspective which is the success of a process. The wide use of the textbook helps the researchers and the inspectors to decide whether some lessons are effective or not and if the teacher is in progress or not.

1.1.1.2.3.   Evaluation in the educational process

Evaluating the language teaching methods, materials, and the effectiveness in teaching means to know what is we are evaluating, how materials are presented and the way teachers design the courses. The evaluation intends to focus on the learning process as a guide; it does not focus only on the results.

1.1.1.2.4.   Evaluation of process and product

The advantage of the evaluation is the ability to recapture what has been taking place during a programme of instruction. With the information gathered during the programme process, the researchers (teachers or inspectors) are able to say why certain materials are learned better. In other words, if we are to make valid statement about improvement caused by particular type of course in the textbook, evaluation helps us to examine what happened during that course. The discussion of how the learner learns "the process" also leads the researchers to what the learner can do in given time or "the product". The product, as a long term, can be viewed as the success of given course. The

product is not matter of only passing a test and succeeding but also it is the proficiency that the learner acquired through an academic course.

1.1.1.2.5. Evaluation and innovation

The teacher uses the evaluation as a means to obtain knowledge about why learners master the given course better with some particular methods, in the main time they fail or encounter great difficulties when they are exposed to other methods. In the innovation or the planned change researchers evaluate the effectiveness of some new methods or judge old ones. The innovation does not focus on only the methods used in class but also on the curriculum such as in textbooks or smaller scale such as new procedures in the development of listening skills for beginners.

1.1.1.2.6. Evaluation and management

The good management provides good learning environment through taking both continuous and final evaluation into any educational plan. The teacher as a manager needs to take the leadership to guide the learners getting the task done effectively using their own experience with learning strategies. The decisions made during the lesson are as important as tasks prepared for the class activities. The teacher taking the role of leader is more than just guiding the learners but he is also evaluating the course; moreover, teachers need to evaluate and adept themselves to the new changes as well.

1.1.1.2.7. Evaluation and context

Evaluation is a very important element in the education process; thus, there is a great interest for teachers or the inspectors to be fair and valid when they design syllabus or make judgments about the learning process. However neither the evaluators (the teachers) nor the learners live in an isolated environment, and even when it comes to the

academic institutions, researchers can notice that the classroom is influenced by individuals and group interaction. Schools, universities and any academic institution are located in a large social and political environment which affects the evaluation in one way or another.

The different elements mentioned before have a direct effect on the validity of an evaluation, starting from the educational system, the changes that might enhance the learning process, the way of how a curriculum is presented, and the general environment.

1.1.1.3.   Types and purposes of evaluation

Scriven (1991) is the first linguist who introduced the concept of formative and summative evaluation in the literature of evaluation. According to him, formative evaluation requires collecting and sharing information for program improvement. It is usually under taken during the progress or improvement of a program, product or person and so on. This type of evaluation serves to:

- Validate or ensure that the objectives of the teacher/instructions are being achieved.

- Provide data to enable on the-spot changes to be made where necessary.

Summative evaluation; on the other hand, is conducted at the end of a program for the sake of:

- Providing data from which decisions can be made.

- Providing useful information and analysis on the products efficacy. For example, finding out whether the learners have learnt what they were expected to learn after using the instructional module.

In fact, both of these two types of evaluation are essential. Regardless to all the differences between them, when we make decisions we need both of them during the development stages of a program to improve and strengthen it and again when it has stabilized to judge its final worth or determine its future (Tunç, 2010, p.20).

1.1.2. Determine the concept of Assessment

The second term that we need to deal with is assessment which is for some linguists a part of evaluation.

1.1.2.1.   Definition Assessment

Assessment is about using evidence whether children have achieved a learning objective. It relies on the ability to make professional judgments using pedagogical and subject knowledge. According to Longman Dictionary of Language Teaching and Applied Linguistics (2002), the term assessment refers to a systematic approach to collecting information and making inferences about the ability of a student or the quality or success of a teaching course on the basis of various sources of evidence. It may be done by test, interview, questionnaire, etc. (as cited in Hanauer, Hatfull and Jackobs-Sera (2009, p. 23)).Zachos (2004) States that assessment "is the process of obtaining evidence to support inferences concerning the attainment of learning objectives" (p. 748). Assessment can be defined as a form of scientific inquiry process in which assessors collect data in order to understand student learning. Assessment can take one of the following forms: presentations, scenarios, projects, portfolios or traditional tests.

1.1.2.2.   Types of assessment

Assessment can be divided to many types. We will discuss the most important ones which we often use.

1.1.2.2.1. Formative assessment

As cited in Higgins, Grant, Thomson and Montarzino (2010, p. 6) "Formative assessment is the short term collection and use of evidence to guide learning". It takes place during the course in a formal or informal ways to provide feedback to inform further development. Even though formative assessment is more concerned with what student carried out during the module, it also may or may not contribute to the final mark. In simple words, formative assessment is a judgment about the learner's performance against a specific objective based on evidence gathered by means of observation, questioning and discussion, learner's responses to classroom activities, and learner' self-assessment. It is used during the whole year to inform future planning.

1.1.2.2.2. Summative assessment

It is a summary of development in a particular time. Typically it occurs at the end of a term or at the end of the school year by means of assessment tasks and tests. Unlike the formative assessment which gathers feedback to enhance the development of learning; thus, it is assessment for learning. However, summative assessment is more concerned with final result and it is known as a summary judgment or final result of the learner's attainments over a period of time, so it is viewed as assessment of learning.

Even though the two concepts have different definitions and uses, it is not possible to use one without the other. If the formative assessment was a journey, the summative assessment is the arrival point. Whereas some researcher claimed that there is no different between the two as it was cited in Higgins, Grant,Thomson and Montarzino (2010, p. 6) "Some educational researchers say that the stark distinction between formative and summative assessment does not really exist" (Ramsden 1992, p. 212-213).

1.1.2.2.3. Diagnostic assessment

Like formative assessment, diagnostic assessment is intended to improve the learners' experience and their level of achievement. However, diagnostic assessment looks backwards rather than forwards; it is also known as "pre-assessments". Diagnostic assessment allows the teachers or the educators to evaluate their student's learning through tests and measurements. However, the main concern of the diagnostic assessment is not measuring but is to ascertain, before instruction, each student' strengths, weaknesses, knowledge, and skills. Establishing these permits the instructor to remediate students and adjust the curriculum to meet each learner's unique needs. The importance of diagnostic assessment appears in the following points:

- Revealing what students already know and do not know about a subject so that the teacher can design lessons appropriately.

- Exposing misunderstandings and misconceptions in prior knowledge.

1.1.2.2.4. Evaluative assessment

This type of assessment refers to an objective judgment of the effectiveness of teaching based on evidence of attainment in identified aspects of the curriculum by means of observation or students" books analysis of test results. It is used during the whole years as well as at the end of the year to identify whether the curriculum is being taught effectively.

1.1.3. Determine the concept of Testing

Finally the last term in the trilogy, the most spread and used method to gather feedback from students.

1.1.3.1. Definition Testing

A test is a social phenomenon which has involved fulfilling a number of functions in the classroom as well as in the society. Basically, it is a procedure intended to establish the quality, performance or reliability of a learning process. The tests can be written or spoken examination of learner's proficiency or knowledge about a given case of study. Deygers (n.d, p. 9) stated in his work, entitled "An Introduction into (English) Language Testing", that Dochy defines a test as a formalized collection of tasks designed to determine a test takers", (people who are going to be tested), ability, knowledge, and/or intelligence. In short, testing is finding out how well a given learning subject is progressing and telling what level of knowledge or skill has been acquired.

1.1.3.2. Characteristics of tests

Anything which can be defined can be tested; you cannot test what you have not taught. Testing may sound for an inexperienced teacher as a simple procedure where he/she puts together some questions and problems to solve; however, testing can be more complicated. Good tests have some characteristics to follow. Bachman suggests in his book "Assessment" (1992, p. 119), that a language test can be classified in terms of five characteristics, which are as follows:

• Tests can be distinguished according to their intended use, such as selection, entrance, readiness, placement, diagnosis, progress, attainment and mastery.

• Tests can differ in content; Achievement tests are based on syllabus, while a proficiency test derives from theory of language ability.

- Different frames of reference can provide the basis for test development and score interpretation. Norm referenced tests are developed to maximize differences among individual test takers and a test score is interpreted in relation to the score of the test takers.

- Tests can be classified according to the scoring procedure (the act or process of evaluating responses to test situations or evaluating characteristics of whomever or whatever is being rated. It consists of checking the student's response to each item to see if it is correct. Scoring objective tests is purely mechanical process which requires no special skills); objective tests require no judgment on the part of the scorer but in subjective tests. The scorer must judge the correctness of the test taker's response.

- Tests may employ different testing methods, such as dictation, cloze, multiple choice, completion, composition and interview.

1.1.3.3.   Requirements of a good and useful test

The well designed test is good test. The test considered as the chief assessment device which is used to gather data for assessing and evaluating educational and training programs; therefore, its results are very important. Thus, a good and useful test is characterized by certain features which can be termed as "requirements of a test". Motiure, et al., (2012) stated in their work that according to Ingram (1974), a good test must meet or have at least six requirements:

1.1.3.3.1. Validity

The key validity question has always been: does my test measure what I think it does? If it does then the test is valid. Content validity is concerned with the degree to which the test questions fairly represent what the test is intended to measure. For example, in our case, if the test does not provide didactic issues to solve, the test will not be valid.

Furthermore, the validity of a test is also about the scores of people on a test for which construct validity is being sought. To sum up, a test is valid when it measures what testers and/or teachers intend it to measure. In other words, if the test is intended to measure communicative ability in English language, for example, and the score obtained is higher or lower due to the test takers" ability, then the test is valid. Additionally, the term validity refers also to the way in which test results are used to make decisions about test takers" learning.

1.1.3.3.2. Reliability

Although validity and reliability can be viewed as synonyms in term of terminology and even seen as closely inter-related, they are completely different concept. Reliability is also seen as a component of validity, but it is not sufficient to judge the validity of a test. A test can be reliable but not valid, whereas a test cannot valid, but not reliable. There are some factors that affect the reliability of a test and we can mention: (1) length; longer test are more reliable, (2) suitable questions, (3) terminology and the marking schedule. To sum up, reliability in testing means "consistency of measurement". It is normally assumed that test takers should receive the same score, taking into account the errors that are expected in tests, if these latter will be corrected by different testers. For example, if a test (w), taken by student (x), is graded twice by teacher (z), student (x) will receive two identical scores, and if the same test (w) is graded by teacher (z) and teacher (y), student (x) will receive two identical scores. In that case we can say that the test (w) is reliable. Simply, reliability is the degree to which an assessment tool produces stable and consistent results.

1.1.3.3.3. Discrimination

Discrimination is one of the most important requirements which is necessary for a test. This requirement is not concerned with sex or race but with the level. While designing a test, teachers must take in mind the level of test takers; it means that a test must be designed in such way that it can discriminate among the students. If we want to measure the height of the school children, we should use a measuring device which is suitable for the students who are going to be tested.

1.1.3.3.4. Economy

This is considered as a practical criterion. The test should measure what it wants to test and it should also measure in a reasonable time. If that happens, we can say that this test is practical and economical.

1.1.3.3.5. Administrability

This means that the test should be as if it may be given under the conditions that exist and the personnel (person who is conducting the test) who are available. For instance, if a test requires electronic equipments and the service of highly trained technician, then it is not administrable because these facilities are not available in most schools and even most colleges and universities.

1.1.3.3.6. Scorability

It means that the test should be scored with ease so that the user may be able to handle it. In addition to that, there should not be any differences in scoring because the difference will affect the accuracy of the test.

1.2.   The history of language testing

According to Spolsky (2001), if we accept the development of examinations on classical Confucian (doctrine during the Hen Dynasty) as the beginning of formal testing, we have 2000 years of history from which to derive our understanding of the process on which we base our assessment of the current state of the art, i.e. tests and examinations have become progressive more powerful (cited in Behrahi, n.d, p. 2).

Language testing is one of the most important sub-fields within applied linguistics. As a methodology for probing and investigating language ability, it comes from a long and honorable tradition of practical teaching and learning needs. It has evolved and expanded in the past decades. He was Bachman (1999) who investigated language testing from 1978; the test design was based on the idea of isolated testing. The tests were formulated on the theories that view language as consisting of skills and component. In 1979, this theories encountered great criticism especially from those who supported the communicative competence. The researchers of this approach such as Widdowson (1983) and Swain (1980) and creation of discourse, or the situation negotiation of meaning and of language ability as multi-componential and dynamic made a great impact on language testing and changed the idea of isolation. The language testing continues to extend with the appearance of SLA. There were many approaches that affect the methods that is used in language testing.

1.2.1. The Different approaches to language testing

For more understanding of the history of language testing, it could be better to investigate the different trends or movements in the realm of language teaching and testing from the onset of its appearance. Both Spolsky (1978) and Hinofotis (1981) (cited in Behrahi, n.d, p. 4).have pointed out that language testing can be broken into periods, or

trends of development. Hinofotis (1981) has labeled the testing trend into the pre-scientific period, the psychometric or structuralist period, and the integrative sociolinguistic period. Likewise, Brown (1996) has used the term movements instead of period to describe them because they somehow overlap chronologically and co-exist today in different parts of the world (As cited in Behrahi, n.d, p. 4).

1.2.1.1. Essay-Translation approach /Pre-Scientific movement: (before1950s)

As cited in Behrahi (n.d, p. 5), Language testing has its roots in pre-scientific stage in which no special skill or expertise in testing is required. This is characterized by lack of concern for statistical consideration or for such notions as objectivity and relativity (Heaton 1988, Weir 1990, Farhady et al., 1994).In this trend, the only judgment of the test is the subjective judgment of an experienced teacher. He can decide what mark to give after reading an essay or hearing a conversation based on language use ability. This trend is characterized by translation tests developed in the classroom by the teacher; thus, these tests luck objectivity as long as the teacher is the only judge. In Hinofotis article (1981), it is mentioned that this trend ended with the onset of psychometric structuralist movement, but in fact such movements have no end in language teaching and testing because such practices are indubitably going on in parts of the world depending on the needs which specific academic contexts demand.

1.2.1.2. Weaknesses of Essay-Translation approach

According to Brown(1996) who explained in his book "Testing in language programs" Essay-Translation Approach was fully based on translation tests developed in the classroom by the teacher. The teacher was the only one who can evaluate and judge

his/her student; the judgments were made subjectively which made it hard to take these marks for granted. In addition, the tests have a heavy literary and cultural bias; thus, the results of a given test are only related to a specific classroom with a specific teacher. To sum up, there are two main weaknesses in this approach:

- Subjective judgment of teachers tends to be biased.

- As mentioned, the tests have a heavy literary and cultural bias.

1.2.1.2. Psychometric Structuralist movements (1950s-1960s)

According to Carrol (1972), psychometric-structuralist tests typically set out to measure the discrete structural elements of language being taught in audio-lingual and related teaching methods of the time. (Cited in Behrahin.d, p. 5). With this movement, language tests became more scientific, reliable; and more precise. In this era, unlike the translation approach that was based on subjectivity, testers and psychologists, who were responsible for the development of modern theories and techniques of educational measurement, were trying to provide objective measure using various statistical techniques to assure reliability and certain kind of validity. The tests in this approach were constructed to be reliable and valid because of the ability to be administrated and the objective scoring; thus, this approach was considered as step forward and improvement on testing practices of the pre-scientific movement. This movement had a great important because for the first time language test development followed scientific principles that make language testing easy to construct, to handle and to score.

1.2.1.2.1. Weaknesses of structuralism approach

According to Brown (1996),Even though this approach was a great improvement in administrating, scoring, and handling; however, it has some weaknesses. The teachers in

this approach encounter great difficulty in constructing the test; in addition, this approach does not focus enough on the integrated skills. To sum up this approach has two major weaknesses:

- It tends to be a complicated job for teachers to prepare questionnaires using this approach.

- This approach considers measuring non-integrated skills more than integrated skills (listening, speaking, reading and writing).

1.2.1.3. Integrative Sociolinguistic movement: (after 1960s)

This movement has certainly its roots in the argument that language is creative, beginning with work of sociolinguists like Hymes (1967). With the English language being global language, professionals viewed that that language is more than the sum of the discrete elements being tested; it goes beyond to the integration of all linguistic abilities. With this new nature of language which characterized globalism, it cannot be captured in additive tests of grammar, reading, vocabulary, and other discrete points of language. The tests used in this movement emphasized on testing language in use through cloze test and dictation, both of which assess the students' ability to manipulate language within a context of extended text. The possibility of testing language in use helped to prove that linguistic elements of language are interrelated and relevant to human experience. Evaluation within a communicative approach must necessarily address, for example, new content areas such as sociolinguistic appropriateness rules, new testing formats to permit and encourage creative, open-ended language use, new test administration procedures to emphasize interpersonal interaction in authentic situations, and new scoring procedures of a manual and judgmental nature (Canale 1984, p. 79; cited in Bachman, 1995).

1.2.1.3.1. Weakness of integrative approach

Bachman (1995) pointed out that this movement was taking the opposite direction of psychometric-structuralism. Whereas psychometric-structuralism emphasized on non-integrated skills, Integrative Approach viewed the integrated skills are what matters. Even if many think that measuring integrated skills is better, sometimes there is a need to consider the importance of measuring skills based on students' needs such as writing only, speaking only, etc. to sum up, the main weaknesses of this approach were:

- Emphasizing on integrated skills

- Neglecting the non-integrated skills

1.2.1.4.   The communicative approach

According to Oller (1986) and Bachman (1995) the attempt to measure different language skills in communicative tests is based on a view of language referred to as divisibility hypothesis. Language use is often emphasized to the exclusion of language usage. Communicative tests are concerned primarily with how language is used in communication and the test content should totally be relevant for a particular group of examines and the tasks set should relate to real-life situation. Communicative testing introduces the concept of qualitative modes of assessment in preference to quantitative modes of assessment.

1.2.1.4.1. Weaknesses of communicative approach

Unlike the structuralist approach, this approach does not emphasize learning structural grammar, yet it may be difficult to achieve communicative competence

without a considerable mastery of the grammar of a language. It is possible for cultural

bias to affect the reliability of the tests being administered (Bachman, 1995). To sum up:

- Communicative approach does not emphasize on learning structural grammar.

- The reliability of the test can be affected by cultural bias.

Conclusion

This chapter has involved an attempt to review the key concepts of language

testing as defined and discussed by many linguists. This interest in the language testing

shows that the terms assessment, testing, and evaluation are different concepts which have

an important part in foreign language learning and teaching. It is also shown throughout

the chapter that teachers should be aware of the most significant relationships of evaluation

with others concepts such as education, management, the context, and innovation. In

addition to that, the instructor should understand what type of evaluation he should use and

for what purpose. This chapter consists also of a detailed discussion of assessment's types.

Teachers need to understand how a test should be .Therefore; the current chapter explained

the most important characteristics that a test should contain. Then, the major requirements

of useful test are discussed through this section. The last point is mainly centered in

showing the different approaches to language testing and the weaknesses of each approach.

Chapter two: Designing Useful Tests

Introduction

Creating a test is one of the most challenging tasks confronting an instructor. Even though, a notable concern of many teachers is that they frequently have the task of constructing tests but they have a little training to rely on in this task. Many views and theories were made about determining the quality of a test, generally most of them have agreed on some areas such as: cognitive complexity, content quality, meaningfulness, language appropriateness, transfer and generalizability, fairness, and reliability and so on.

Teachers most of the time use tests that they have prepared by themselves based on their experience. However, the lack of experience may lead to undesirable effects such as tricky questions which sometimes take place unintentionally. Using the inappropriate type of test is very common for new teachers. Well-constructed tests motivate students and help teachers to make a better assessment about the learning process. A simple question as how can we have a good test can be answered by "does the test measure what it should measure" for this sake, the constructor of the test must take in consideration several points: the appropriate length; the more items the tests have the more reliable they are, and for the students, the more questions there are the less effective the wrong answers will be. To avoid the misunderstanding of question, it is crustal to   give clear instructions in the test; in addition the test should include all types of questions (multiple choices, true/false, essay). If possible, a little use of humor can increase the sense of relaxation which can help student to deal with the difficult questions (Clay 2001).

The formal training for test instructors can be very helpful, however; some studies proved that experience is the most important factor in this field. In this chapter we will discuss the most important element that can influence the quality of designing tests for

both experienced and non-experienced teachers. This part of our work will also deal with the test as a means of learning.

2.1.   Characteristics of usefultests

Thompson (Forum, 2001) believes that students learn more when they have tests. This saying can be true to some extent; basically students will prepare only the materials that they will be tested in. Even if students master these materials which mean more than 75% of it is perfectly learned this process still called parietal learning because other materials have been ignored. Concerning the level of students, it is very important to prepare tests that suit all students otherwise the feedback is not reliable. To sum up we can add alluding to Alderson (1996) that the usual classroom test should not be too complicated and should not discriminate between the levels of the students. The test should test what was taught thoroughly. For this sake, there are some characteristics that make tests useful and reliable. Clay (2001) provided the criteria for establishing technical quality of a test. In addition to reliability, he suggested also:

2.1.1. Cognitive complexity

Cognitive complexity refers to the various levels of learning that can be tested. The test should reflect the goals of the instructor; if the instructor is concerned with students' memorization, the test should contain simple recalling questions. If the instructor is concerned with the students' analytic abilities, the test should ask analytic questions. The levels of learning refers to the various levels of learning that can be tested. These levels are known as Bloom's Taxonomies of Educational Objectives. He listed six levels of intellectual understanding which are: knowledge, analysis, comprehension, application, synthesis, and evaluation.

2.1.1.1. Bloom's Taxonomies

During the 1948 convention of American psychological association, a group of educational psychologist decided that it would be useful to classify the levels of understanding. In 1956, Dr. Benjamin S. Bloom published his work "Taxonomies of Educational Objectives" that lists the six levels.

2.1.1.1.1. Knowledge

This level is based on the simple questions (who, what, where…) in order to recognize and recall information (date, event, definition…).

2.1.1.1.2. Comprehension

The comprehension level is based on the student's understanding, the meaning of information. In this level, the teachers use questions such as: Restate in your own words, List three reasons for, translating from one form to another. This level is usually about interpreting, explaining, and summarizing.

2.1.1.1.3. Application

In this level the students have the rules, methods or the principles; all they have to do is apply them. The instructor usually asks students to classify something or use rules to solve a problem.

2.1.1.1.4. Analysis

The students in this level can identify the component within a system(phrase, paragraph,...) and also they can find the relationship between the components.

2.1.1.1.5. Synthesis

The students in this level are expected to be able to discover and create new connections by combining ideas then develop new ideas of their own. The questions that

are used in this level focus on what the students can add or develop (What would you infer from...?, What ideas can you add to... ?...)

2.1.1.1.6. Evaluation

As a highest level of learning, the students can use evidence and arguments to judge the quality of a test through resolving controversies and differences of a given opinion. The teacher uses certain type of questions such as: How would you decide about…?, What priority would you give…?.

These are the levels of understanding. The teachers should take them in consideration when planning a test to give more reliability to the feedback.

2.1.2. Content quality

The quality of the content is a very important element; the questions of the test will not only challenge the students' knowledge about a subject but also will permit them to demonstrate their knowledge (Davis and Barbara 1999).

The first step in constructing a test is planning, and within the planning, the instructor should first prepare an outline about what the test will cover. The teacher must take in consideration the subjects he/she emphasized on in the classroom plus the important concepts that were covered in each class if there is more than one class(i.e. if the teacher focused on Shakespeare and discussed mutual concepts in all classes then he/she needs to make at least 60% of the test content about this subject) (Clay, 2001). If the test was a less focused subject will be regarded as unfair to the students because they have naturally focused on Shakespeare while revising.

Concerning the content quality, there are some questions which need to be answered: What skills do they indicate to be tested? How many questions and how many areas will be covered? How many sections will there be in the test? What are the

specifications? And finally, what formats will be used to test. It is important to explain the specifications and the format of the test thoroughly (Clay, 2001).

2.1.2.1.   Test specification

According to Alderson, Clapham, and Wall (1995), test specifications are also called "blueprints" or design documents of a test. They tell test writers, designers, and test assemblers how to write items and how to construct new forms of tests. Test specification can take many forms. It could be a single document that describes the test purpose, test constructs, the one or two item types that appear on the test, and statement of the number of items that the test as a whole should contain. Alternatively, it may be necessary to have a separate document for each item type. The test is complex, the more specification documents there are likely to be (cited in Glenn, 2010). It is possible to list all the possible specifications that we might need for a test, following the plan of Mislevy, Almond and Lukas (2003):

2.1.2.1.1. Item/task specifications

Item/ task specifications are the most important part in test specification. They describe the prompts that are designed to elicit the evidence upon which the inferences are made about the targeted abilities of the learners. These specifications are concerned with the kind of materials that test takers will encounter, they also should state what the instructions look like(mostly, instructions are introduced with questions but sometime they came as a not in the end or as oral explanaition) in addition to any other features that are important.

2.1.2.1.2. Evidence specification

It refers to what the test taker is expected to do or what kind of response is expected in each task. It may be found in some references as "response attribute". It also

states which response is to be scored. This is considered as the measurement component for a closed-response item such as multiple choices activities. This may simply be that each response is scored as 0 for incorrect answer and 1 for the correct one; however, for performance tasks, it would be necessary to provide more complex rating scales or other devices to guide the judgments of human assessors.

2.1.2.1.3. Test assembly specification

This document provides the instructions for how the entire test is constructed. Test designers might know that there are four item types, but they still need to know how many of each item type they need in the test. They also know that reliability in tests is directly related to the test length; therefore, they may need to specify the target reliability and the minimum number of items needed to meet the target. The test assembly specification therefore plays a critical role in showing that the number and range of items in any form of the test adequately represent the key features of the criterion situation in the real world.

2.1.2.1.4. Presentation specification

This specification tells the test designers how the items and any supporting material are to be presented to the test takers.

2.1.2.1.5. Delivery specification

This document sets out the detailed of the test administration, test security, and test timing. This specification may include spacing between desks or computers, the number of invigilators (who watch over students in an examination) required per number test, what may or may not be used during the test (dictionaries, handouts, etc…), how long each sub-test should take, and what the overall time allocated to the test is. All of these specifications serve to construct and deliver "test formats" for particular administration. A test is really an abstract term, i.e., a test is really a collection of specifications and any

realization of specifications a test form. The first step in planning a test is to outline the actual course content that the test will cover by taking few minutes following each class to register on an index card of the important concepts covered in class and in assigned reading for that day. These cards can be used later as a source of a test items.

These specifications will increase the quality of the test; the well prepared tests help the test takers to feel at ease during the test simply because the test is well formed, the instructions are clear and the time is sufficient.

2.1.2.2. Test items formats

The quality of a test is influenced by the format of the test; it is very useful to apply different formats in the same test.

2.1.2.2.1. Multiple choice test items

Multiple choice items are considered to be among the most common item types. It can be used to test factual recall as well as students' understanding.

> Understand that there is always one clearly best answer. My goal is not to trick students or require them to make difficult judgments about two options that are nearly equally correct. My goal is to design questions that students who understand will answer correctly and students who do not understand will answer incorrectly. Johnson.

The multiple choice item consists of the stem, which identifies the question or problem and the response alternatives or choices. Multiple choice items are considered to be among the most versatile of all item types. They can be used to test factual recall as well as levels of understanding and ability to apply learning.

2.1.2.2.2.   Multiple choice test items types

The multiple choice test can take three different types

• Question/ right answer; in this type the students/test takers need to chose the correct answer among different options that are given by the instructor.

• Incomplete statement; usually this type used in definitions or famous sayings.

• Best answer; this type depends partially on the student understanding to identify the best answers.

2.1.2.2.3.   Advantages and disadvantages of multiple choice test item

The multiple choice test has great deal of advantages, it is very effective and Versatile at all levels. The multiple choice test can cover the whole program in the minimum time that is given in the test because there is a little writing is done by the students and the guesses are reduced. Even though the multiple choice test has many advantages, it has some disadvantages. It is difficult to construct good test, and it is also difficult to come up with plausible distracters / alternative responses.

2.1.2.2.4.   Some tips for the instructors to design good multiple choice test item

• Use at least four alternatives for each item to lower the probability of getting the item correct by guessing.

• Avoid irrelevant clues such as grammatical structure, well known verbal associations or simplistic connections between stem and answer.

• In testing for definitions, use the term in the stem rather than as an option.

• Keep all alternatives in a similar format.

2.1.2.2.2.True-False test items

The main idea of the true/false is the forced choice because the students must choose between only two possible answers.

As cited in Clay (2001) a major distinction between the true-false test item and items in a multiple choice format is that the true-false statement contains no criterion for answering the question. Each examinee must ask the question: True or false with respect to what? Each true-false item must be unequivocally true or unequivocally false. It is imperative that proper wording and the elimination of extraneous clues are more crucial with the true-false item than with any other test format. "Writing Test Items"

In the most basic format, true-false questions are those in which a statement is presented and the student indicates in some manner whether the statement is true or false. It is different from the multiple choice test because the students have only two options true/false. This format of test based on students' recognition of the correct and the incorrect answers. The true/false test can be extremely easy or extremely difficult because sometimes students take chances and rely only on their luck.

2.1.2.2.2.1. Advantages and disadvantages of the true/false item

The true/false tests are used to evaluate students' understanding of popular misconceptions and they can take less time in constructing or answering them. For the disadvantages, they are known among students as "luck test" because all answers have 50/50 chance, and it also needs a large number of items for high reliability (Clay, 2001).

2.1.2.2.2.2. Some tips for the instructors to design good true/false test item

- Keep the language as simple and clear as possible.

- Be aware that extremely long or complicated statements will test reading skill rather than content knowledge.

- Randomize the sequence of true and false statements.

- Write items so that the incorrect response is more plausible or attractive to those without the specialized knowledge being tested.

2.1.2.2.3. Matching test items

A simple matching item consists of two columns: one column of stems or problems to be answered, and another column of responses from which the answers are to be chosen. Matching questions provide a most efficient way to test knowledge in courses in which events, dates, names, and places are important. Matching tests can take different types such as: Terms with definitions, Phrases with other phrases, Causes with effects, Parts with larger units, and Problems with solutions.

2.1.2.2.3.1. Advantages and disadvantages of the matching test item

The matching tests cover maximum amount of knowledge in minimum space and it is very valuable in content areas that have a lot of facts. However, it is time consuming for student plus, it might not be appropriate for high levels of learning.

2.1.2.2.3.2. Some tips for the instructors to design good matching test item

- Review your teaching objectives to make sure that a matching component is appropriate.

- When possible, reduce the amount of reading time by including only short phrases or single words in the response list.

- Make sure that there are never multiple correct responses for one stem.

2.1.2.2.4. Fill-in-the-Blank test items

Completion items are especially useful in assessing mastery of factual information when a specific word or phrase is important to know. A completion item requires the student to answer a question or to finish an incomplete statement by filling in a blank with the correct word or phrase. Completion items tend to test only rote, repetitive responses and may encourage a fragmented study style since memorization of bits and pieces will result in higher scores. They are more difficult to score than forced-choice items and scoring often must be done by the test writer since more than one answer may have to be considered correct.

2.1.2.2.4.1. Advantages and disadvantages of Fill-in-the-Blank test items

This test item is good when the content is related to who, when, what and where questions. It also encourages the intensive study because the student must know the answer plus, he/she must know where to put it. However, this item has some disadvantages; in addition to that, one question may have more than one answer. This test item overemphasis the memorization and it is time consuming for students.

2.1.2.2.4.2. Some tips for the instructors to design good Fill-in-the-Blank test item

- Avoid using a long quote with multiple blanks to complete

- When working with definitions, supply the term, not the definition for a better judge of student knowledge.

2.1.2.2.5. Essay test items

A typical essay test usually consists of a small number of questions to which the student is expected to recall and organize knowledge in logical, integrated answers. The main advantages of essay and short answer items are that they permit students to

demonstrate achievement of such higher level objectives as analyzing and critical thinking. There are two main types in essay item test: (1) Extended response in which the students are allowed to write freely with little instructions provided by the teacher. (2) Restricted response; in this type the students are supposed to follow the outline that is provided by the teacher.

2.1.2.2.5.1.   Advantages and disadvantages essay test items

The students in this item are less likely to guess the answers because it is most of the time indirectly provided. This item also allows students to demonstrate ability to organize knowledge, express opinions, and show originality. From the other side, this item is much known with time consuming and it is also subjective and potentially unreliable in term of scoring.

2.1.2.2.5.2.   Some tips for the instructors essay test items

•   Make essay questions comprehensive rather than focused on particular units of content.

•   Allow students an appropriate amount of time.

•   Inform students, in advance of answering the questions, of the proportional value of each item in comparison to the total grade.

•   Students should be informed about how you treat such things as misspelled words, neatness, handwriting, grammar. The teacher should respect these instructions as well.

2.1.3.   Meaningfulness

According to Dewey, PhD(Georgia Southern University, Statesboro, GA), instead of forcing students to guess what will be included in the test or to psych-out the teacher to decide what to study, it is better to give students specific study questions, then draw the test from the study questions. Sometimes this sounds the same as teaching the test. This

may be true if there are very few study questions, but if the teacher offers questions for all of the most important ideas in an assignment, then teaching the test is teaching the course (cited in Clay, 2001).

2.1.4. Language appropriateness

When constructing or preparing tests items, the language being used should be simple and clear as it is in the classroom. They should be free of non-functional material and extraneous (not directly connected with what is being dealt with) clues. They should also be free of race, ethnic, and sex bias. Beyond these two qualifications, students' language back-grounds impact their performance on tests. Modifications of the test for students that are limited English proficient include: assessment in the native language, text changes in vocabulary, modification of linguistic complexity, addition of visual supports, use of glossaries in native language, use of glossaries in English, linguistic modification of test directions, and additional example items/tasks (Clay, 2001).

2.1.5. Transfer and Generalizability

Unlike the traditional testing, all the other forms of assessment (presentations, scenarios, projects, and portfolios)add dimensions to the process of assessment. Teachers can make valid generalizations about achievement more easily by using authentic and performance assessments. These generalizations may involve instructional placement decisions. Well-structured tests allow teachers to understand what needs to be taught next. Teachers are also able to monitor a student's learning and can change the instruction program as needed.(Clay, 2001).

2.1.6. Fairness

The students' performance should not be measured in a way that gives advantage to factors that are irrelevant to school. The main objective of test providers is to make their tests as fair as possible. Council of Europe (2011), acknowledged three aspects of fairness:

- Fairness as lack of bias.

- Fairness as equitable treatment in the testing process.

- Fairness as equality in outcome of testing.

Grading constructively requires the instructor to provide feedback (written and/or oral) that helps the students to appreciate their achievement in the test Clay (2001) suggested few basic rules of fairness which are:

- Test questions should reflect the objectives of the unit.

- Expectations should be clearly known by the students.

- Each test item should be presented in a clear formulated task.

- Enough time for test completion should be allowed.

2.1.7. Reliability

The main reason of the test is to enhance learning. The good test is the test that is designed with items that are not easily guessed or answered without preparing, even though; it is possible to construct a test that is fully based on comprehension. Answers to test questions will be consistently trusted to represent what students know.

A very important point became an issue in the reliability aspect that is multiple-choices test versus essay test; in multiple-choice test students are very likely to guess

correctly just by knowing about the subject matter (have interest in the subject even if they did not prepare). The instructor when planning for this type of test should follow some rules such as:

- Make sure each answer is used the same number of times, in random order.

- Make sure such answers are correct about a fifth of the time.

- Sometimes offer opposites when neither is correct.

- Use scientific sounding jargon in wrong answers.

- Sometimes make the simple, obvious answer the correct one.

On the other hand, the essay test is known with that grades given to the answers are mostly depending on neatness and handwriting. It is also hard to ensure that the work is original. The students in this type of test attempt to write lengthy essay in the sake of increasing the chances to meet with the teachers' desired answers. The essay test requires limitations that are clearly stated in the test and highlighted points that the students should discuss.

2.2. The different types of tests

Unlike what most students feel about the tests, the teachers do not attempt to punish the unprepared students, it is also not the motivating factor for the students to study even though that is partially needed. In fact, the test is a request for information and possibility to learn what the teachers did not know about their students before. We can add here that the test is important for the students, too, though they are unaware of that. The test is supposed to display not only the students' weak points, but also their strong sides. According to some scholars (Thompson, 2001; Hughes, 1989; Alderson, 1996; Heaton,

1990; Underhill, 1991), there are four traditional categories or types of tests: proficiency

tests, achievement tests, diagnostic tests, and placement tests. However, there are two sorts

or kinds of typology method and purpose.Motiure, et al., (2012)

2.2.1. In Terms of method

A broad distinction can be made between pen-and-paper language tests and

performance tests. In performance tests, language skills are assessed in an act of

communication, e.g. tests of

speaking and writing:

Extended samples of speech/writing are elicited.

- Judged by trained markers.

- Common rating procedureused.

Paper-and-pen tests are typically used for the assessment of:

- Separate components of language (grammar, vocabulary …).

- Receptive understanding (listening & reading comprehension).

2.2.2.   In Terms of purpose

In addition to Thompson(2001), Hughes (1989), Alderson(1996), Heaton(1990)

and Underhill(1991)many other scholars distinguished:

2.2.2.1. Achievement/Attainment tests

Alderson (1996) claimed that achievement tests are "more formal", whereas

Hughes (1989, p 8) assumes that this type of tests will fully involve teachers, for they will

be responsible for the preparation of such tests and giving them to the learners. We can

differentiate between two kinds of achievement tests: final and progress tests. Final tests

are the tests that are usually given at the end of the course in order to check the students'

achievement   and   whether   the   objectives   have   been   reached.   To   sum   up;

Achievement/Attainment test is usually more formal. It is designed to show mastery of a

particular syllabus (e.g. end-of-year tests, school-leaving exams, public tests) though similar (re-syllabus) to progress tests; it is rarely constructed by classroom teacher for a particular class, and it is designed primarily to measure individual progress rather than as a means of motivating or reinforcing language.

2.2.2.2. Progress tests

According to Alderson (1996), progress test will show the teacher whether the students have learnt the recently taught material successfully. Most classroom tests are progress tests. They are used to assess the progress students make in mastering materials taught in the classroom; they are often given to motivate students. In addition to the assessment of the degree of success of teaching and learning, they are also used to identify areas of weakness and difficulty. Progress tests can also be diagnostic to some degree. Basically it will display the activities based on the material the teacher is determined to check. In the sake of evaluation, the teacher can develop a system of points that will build up marks. Typically, such tests do not influence the students' final mark at the end of the year. However, if the majority of the class scores are low, the teacher should be cautious. This could be a signal that there is either something wrong with the teaching or the students are low motivated (Ozerova, 2004).

2.2.2.3. Diagnostic tests

Diagnostic Tests are primarily designed to assess students' knowledge & skills in particular areas before a course of study is begun, i.e. it is typically the first step each teacher, even non-language teacher, takes at the beginning of a new school year. According to Longman Dictionary of LTAL (106) diagnostic test is a test that is meant to display what the student knows and what s/he does not know. Hughes (1989, p. 6) adds that diagnostic tests are supposed to spot the students" weak and strong points. Heaton (1990, p.13) compares such type of test with a diagnosis of a patient, and the teacher with a doctor

who states the diagnosis. Underhill (1991, p.14) adds that a diagnostic test provides the student with a variety of language elements, which will help the teacher to determine what the student knows or does not know (cited in Ozerova, 2004, p.13). We believe that the teacher will intentionally include the material that either is presumed to be taught by a syllabus or could be a starting point for a course without the knowledge of which the further work is not possible. Thus, we fully agree with the Heaton's comparison where he contrasts the test with a patient's diagnosis.

2.2.2.4. Placement tests

These tests are used to assign students to classes/programs appropriate to their level of proficiency and to define characteristics of each level of proficiency. According to Longman Dictionary of LTAL (279-280), placement test is a test that places the students at an appropriate level in a program or a course. This term does not refer to the system and construction of the test, but to its usage purpose. According to Hughes (1989, p.7), this type of test is also used to decide which group or class the learner could be joined to (cited in Ozerova, 2004). This statement is entirely supported by another scholar, such as Alderson (1996, 216), who declares that this type of test is meant for showing the teacher the students' level of the language (cited in Ozerova, 2004).

2.2.2.5. Proficiency tests

Regarding Longman Dictionary of LTAL (292) proficiency test is a test, which measures how much of a language a person knows or has learnt. It is not bound to any curriculum or syllabus, but is intended to check the learners' language competence Hughes (1989, p.10) gives the similar definition of proficiency tests stressing that training is not the thing that is emphasized, but the language. He adds that 'proficient' in the case of proficiency tests means possessing a certain ability of using the language according to an

appropriate purpose. We can say, these tests are used to measure students' achievements in relation to a specific task which they are later required to perform (e.g. follow a university course in the English medium; do a particular job). Reference forward to particular application of language acquired: future performance rather than past achievement. They rarely take into account the syllabus that students have followed: Definition of operational needs, Practical situations, authentic strategies for coping, common standard (e.g. driving test regardless of previous learning), and application of common standard whether the syllabus is known or unknown.

2.2.2.6. Aptitude tests

Aptitude tests measure students' probable performance. Reference forward but can be distinguished from proficiency tests. Aptitude tests assess proficiency in language for language use (e.g. will S experience difficulty in identifying sounds or the grammatical structure of a new language?) while Proficiency tests measure adequacy of control in L2 for studying other things through the medium of that language.

2.3. The Stages of test preparation

There are three main stages of preparation of a test:

2.3.1. Planning

This stage covers outlining test, listing of topics, and casting of ideas for items and material collection. In this stage, test developers should take in consideration the characteristics of test takers(their age, gender, social situation, educational situation, mother tongue, etc.), the purpose of the test whether it is for(school-leaving certificate, admition for a program of education, minimum professional requirements, formative or diagnostic function, etc.),the relationship between the test and educational context, the standards that should be followed for the proposed purpose and how the results should be used.

2.3.2.   Composition

This includes the composition of actual items and choice for objectivity. In other words, the aim of this stage is to provide live test materials that are made according to the test specifications and that are ready on time. The process of assembling tests breaks down into three broad stages: Producing materials, quality control, and constructing tests.

2.3.3.   Analysis

The instructor needs to have predictions about the students' level and the appropriate methods that give the test reliability and fairness. The analysis consists of determining difficulties and discrimination of test items, speediness of test, and scope for its improvement.

2.4. Language testing and teaching

According to Hicks (2000, p.155), a test has a great role, especially in language learning. It is a means to show both students and teachers how much the learners have learnt during a course. This is obviously true, we believe that in order to see whether the students have acquired the material and are making constant progress, the teacher will inevitably have to test his/her learners. Heaton (1990, p.6) stated that tests could be used to display the strength and weaknesses of the teaching process and help the teacher to improve it. They can demonstrate what should be paid more attention to, should be worked on and practiced. Furthermore, the tests' results will display the students their weak points, and if carefully guided by the teacher, the students will be even able to take any remedial actions (cited in Ozerova, 2004, p.5). In addition to these purposes, we can say that tests serve to motivate students to learn and to provide a record for assigning grades. Even though the student has to study the material that is supposed to be tested, but often it does not mean that such type of learning will obligatory lead to acquisition and full understanding of it. However, the students' success in the test will motivate them to study

more, and encourage them to proceed even if it is sometimes difficult. Hence, we can speak about assessment and evaluation as means for increasing the students' motivation.

2.4.1. Ways of testing

2.4.1.1. Direct versus indirect testing

according to Hughes (1989, p.14) direct testing means the involvement of a skill that is supposed to be tested; that means when applying the direct testing, the teacher will be interested in testing a particular skill, e.g. if the aim of the test is to check listening comprehension, the students will be given a test that will check their listening skills, such as listening to the tape and doing the accompanying tasks. Such type of tests will not engage testing of other skills. The advantage of the direct test is that focusing on some abilities, so preparation for that usually involves persistent practice of certain skills.

In other hand, Indirect testing, regarding to Hughes, tests the usage of the language in real-life situation. Moreover, it suits all situations; whereas direct testing is bound to certain tasks intended to check a certain skill. He assumes that indirect testing is more effective than direct one, for it covers a broader part of the language.

To sum up, direct tests require the test takers to perform precisely the skills that the test wishes to measure. It is easier to carry out when it is intended to measure speaking and writing skills as well as it has a number of attractions. Additionally, it has a good impact on both teaching and writing processes. It can take the form of a writing composition that aims to measure the students' writing skill. In the other hand, indirect tests attempt to measure the abilities that underlie the skills in which the test is interested. It may take the form of underlined items which the student has to identify as erroneous in Standard English language.

2.4.1.2.   Discretepoint versus integrativtesting

According to Longman Dictionary of LTAL (112), discrete point test is a language test that is meant to test a particular language item, e.g. tenses. The basis of that type of tests is that we can test components of the language (grammar, vocabulary, pronunciation, and spelling) and language skills (listening, reading, speaking, and writing) separately, i.e. it refers to the testing of one element at a time, item by item. We can declare that discrete point test is a common test used by the teachers in our schools. It may take the form of a series of items testing a particular grammatical structure; however, Integrative testing requires the test takers to combine many language elements in the completion of a task. It may take the form of comprehension of words and the ability to use them correctly, free compositions, or cloze tests.

2.4.1.3.   Norm-referenced versus criterion referenced testing

According Longman Dictionary of LTAL (17) that states that criterion-referenced test measures the knowledge of the students according to set standards or criteria. This means that there will be certain criteria according to which the students will be assessed. There will be various criteria for different levels of the students' language knowledge. The teacher do not attempt to compare the results of the students, it is just to measure learners' knowledge of the subject.

From the other side, we have norm-referenced test that measures the knowledge of the learner and compares it with the knowledge of another member of his/her group. According to Hughes (1989), this type of test does not show us what exactly the student knows. Therefore, it is less reliable in term of students' general acquisition, but it can be helpful as a placement test.

2.4.1.4. Objective versus subjective testing

According to Clay (2001), Objective testing requires students to select the correct answer from several alternatives or to supply a word or short phrase to answer a question or complete a statement. It includes multiple choice, true-false, matching, and/or completion tasks whereas; Subjective Testing is based on an opinion or judgment, of the tester, which is expected to match with that of a test taker. It involves more of memorization on the part of the learner. It includes a short answer essay, extended-response essay, problem solving, and/or performance test items.

2.4.2.   The impact of testing on teaching and learning

The impact of testing on teaching and learning in the classroom is known as "Washback or Backwash". Many research studies reveal that a test affects participants, processes and products in teaching and learning. Students, classroom teachers, administrators, material developers and text book writers may be included under "participants". Their perceptions and attitudes towards their work are likely to be affected by a test. Alderson and Wall (1993) stated that a test will influence what and how teachers teach, what and how learners learn, the rate and sequence of teaching and learning, and the attitudes to the content, method, etc. of teaching and learning. It can be positive or negative.

Conclusion

Language testing is one of the most important and complicated tasks that both teachers and learners encounter in their educational process. It has a bad or negative sounding for both because of the hard work that it requires before, during and after the measurement process. Whatever the type of the test is its purpose, its design or method or even the grades it deserves whether they are high scores or low, it is always considered as good tool that gathers data about students' ability to acquire a language and how well they

use it in a good way. The tests can facilitate the students' acquisition process and function as a tool to increase their motivation; however, too much of testing could be disastrous changing entirely the students' attitude towards learning the language, especially if the results are usually dissatisfying.as the test types differ according to the need of the test-maker it also has characteristics that all types of test share this keeps the test useful and trust worth.

Chapter three: Field Work

Introduction

This chapter is devoted to the practical part of our research. It describes in some details how we conducted the tools we used to check whether the TEFL tests in English Department at university of Biskra are applying the basic rules of constructing good and useful tests.

In addition, we will try to verify to what extent providing a descriptive account will be helpful to overcome these problems and promote the tests feedback. For this purpose, we designed a formal questionnaire, interview and we analyzed some TEFL tests that are available to collect data from our sample of interest. We will also provide the analysis to the results obtained and discuss the findings of our study.

3.1. Documents analysis

In this analysis we attempt to much the TEFL tests with what they should contain (the test requirements). We are not in any level attempting to judge teachers.

3.1.1. The first sample

From the sample we gathered we can divide our analysis into two parts form and content.

3.1.1.1. The form

The form is the outlook of the test, the first thing that the student sees. The form of the test can have a positive or negative affect on the student.

The first thing that we have noticed is that teacher (1) pays a great attention to make a well structured test which can attract the test-taker and make him/her feel more at ease. The tests that we have gathered have the same characteristic which is the use of all kinds of test forms (multiple choice, true/false…). No one can deny that these tests are fairly long even though some section are not time consuming but it can be an issue to

decide which part to start with. The issue of the number of the pages which is usually more than 3 can be a shock for the students when first encounter this also can increase the level of anxiety. The point that has a very high attention from the students and which we had noticed is the mark scale; giving very few marks to each question can be frustrating to the focus of the students with questioning which of the exercises he/she should start with. The most important thing that we have noticed is the gradual difficulty in the questions which is most appreciated because the focus of test-taker will not be lost by the up and down in the questions.

3.1.1.2. The content

The content is what the test holds between the line; the questions and the information that need to be analyzed.

We will try in this part to measure the appearance of the most important requirement of tests.

3.1.1.2.1. Validity

The validity means that tests should test what they are meant to test. In the sample that we have the teacher clearly tested the memorization; this can be observed in the matching, true/false and fill-in-the gaps these are slandered items to test the students' memory. Testing understanding and the critical thinking of the student is clear in the last part of the test in which the student is asked to give opinion or to analyze a passage. All these items made the test highly valid.

3.1.1.2.2. Reliability

The reliability means that the results of the tests can be taken as a solid source. Again the use of all test forms and testing all levels of learning can be considered as a reliable source of learning process.

3.1.1.2.3. Suitability

Time, length and score can be gathered to make one big problem which can be called suitability. The student most of the time encounter some difficulties in keeping up with the language because of the too much information that they need to deal with also the ability to keep high focus for the whole one hour and a half can be hard if the test consists pages that covers the whole programme this problem causes low scores.

The long experience combined the high education has made a sourceful tests that can provide a solid information about students and the learning process as well.

3.1.2. The second sample

3.1.2.1.   The form

As former explained the form can have great impact on the student usually it will lead to either to failure because of anxiety or success because of the release.

The first thing that we have notice is the over emphasize on writing. The test is filled with paragraphs that can make the student's concentration easily lost between reading, understanding and trying to form good contributions. Even though the instructions were "write briefly" the students always write as much as they know about that subject which has been proved every time by many researchers. The use of only one test form cannot cover the whole program and the attention of doing that through this test form means that this test will be frustrating to student and time consuming.

3.1.2.2.   The content

We will try in this part to measure the appearance of the most important requirement of tests.

3.1.2.2.1. Validity

Maybe these tests take only one form but that does not make it invalid. The questions contain some memorization items; even though it is meant to be memorization but it has a lot of writing to do which some time can be challenging to keep the mind focus

on all the questions (read, understand then answer) can be a heavy weight to do it for one and a half hour answering long passages. Testing understanding and critical thinking is also viewed in the last part as giving an opinion about a situation or analyzing another passage which can be frustrating. These tests can be taken as a good source to gather information about the writing skill but it is not valid to be taken as a source for the students' knowledge about didactics.

3.1.2.2.2. Reliability

All researchers in the field of tests have agreed on that essay tests are not reliable because they are subjective. Maybe the teacher pays more attention to the content but it is not sufficient because the same content can be influenced by the good or the bad writer.

3.1.2.2.3. suitability

In this sample we also noticed that the content that is supposed to be discussed in the answer sheet is not suitable with the time. The student cannot write brief answers because he/she is interested in meeting with the teachers' request , this can be time consuming and most of the time students' do not finish the test or do not do it appropriately.

3.2.   Analysis of students' questionnaire

After analyzing the test sample we gave away this questionnaire to gather data about the opinion of students about TEFL tests. We have directed this questionnaire to 50 master one English students in random. These students have studied didactics for at least two years and dealt with different didactic teachers.

3.2.1. Section one : General Information

1.   **How can you describe your level in English language?**

Options	Subjects	percentage
Average	17	34%
Good	22	44%
Very good	10	20%
Excellent	1	2%
Total	50	100%

Table one : the level of the students

As the table indicated 44% of the students think that they have a good level in English langue while 34% think that they are average students. This indicate that students are realistic about their level plus it shows that most of them have already made a self-evaluation which is a very important step to be a good learner.

2.   **On what basis do you build your self-evaluation?**

Options	Subjects	percentage
Tests' marks	42	84%
Language understanding		0%
Language use	8	16%
Others		0%
Total	50	100%

Table two : basis of self-evaluation

84% of the students evaluate themselves according to their test marks which indicates that the test marks is the most important element in students' decisions about

their level; because they usually depend on the teachers' opinion about them, it is also their only way to pass from one year to the next.

3.   What do tests mean to you?

Options	Subjects	percentage
Marks and success	37	74%
Identifyweaknesses and strengths	11	22%
Contest to compete	2	4%
A chance to improve skills		0%
Others		%
Total	50	100%

Table three : what the tests means

As we mentioned before test usually means success to most students in our sample 74% view tests as a means of pass from a level to another or as a great achievement that provides a higher level of self-confidence.

4.   Do tests reflect the real level of students?

Options	Subjects	percentage
Yes	23	46%
No	27	54%
Total	50	100%

Table four: the tests and the level of students

5.   Of the following test forms you prefer more? (You can tick more than one).

Option	Subject	percentage
Multiple choice	47	94%

True/false	40	80%
Matching	48	96%
Essay	50	10%
Total	50	100%

Table five: what students prefer

When we analyze those results the first thing we notice is that students do not prefer the essay form of the test. Also the results show us that great number of students prefer the tests that are based on either memorization or general understanding of subjects discussed in the classroom. On other hand we can say that students dislike the analytic part of the test.

6.   Do you appreciate the test to be mixture of all forms?

Options	**Subjects**	**percentage**
Yes	47	94%
No	3	6%
Total	50	100%

Table six : students appreciation for the test mixture

These results show that the vast majority of the students (94%) have a great appreciation to the mixture form of the test. They have also explained that the mixture test give a better chance to score higher.

7.   Do you think that tests are the best way to evaluate and decide whether student succeeded or failed?

Options	Subject	percentage
Yes	15	30%
No	35	70%
Total	50	100%

Table seven : reliability of the test judgment

From this table we can observe that most students (70%) agreed on that tests should not be the judge on the success or the failure of a student. They have explained that there are many factors can affect the student performance in a given test. The test can be a means of measuring the level of the student but it should not stand alone. Some of the student mentioned that the final result of the student level should be based on both summative and formative assessment.

3.2.2. Section Tow: TEFL and tests

1.   **Do you have difficulties in TEFL tests**

Options	Subjects	percentage
Yes	19	38%
No	31	62%
Total	50	100%

Table eight : The difficulty of TEFL tests

62% of the student did not encounter any difficulties in TEFL tests. According to this sample this result is because of their good understanding of TEFL topics. They also explained that the test instructions are often simple and clear.

2.   **If yes what kind of difficulties you often encounter**

Options	Subject	percentage
Time	15	80%
Content	4	20%
Language used	0	0%
Other	0	0%
Total	19	100%

Table nine: the difficulties the students encounter

In the 38% who have difficulties in TEFL test 80% of them had big problem with time. They explained "the length of the test is unbelievable, it needs at least 2 hours to go through it all" (one of the student's explanations). The differences between the students like the time they need to understand or to construct their ideas can an issue because some time what seems to long for some people can be view as too short to others.

3. How often do you feel that TEFL tests are repetitive

Options	Subjects	percentage
Never	0	0%
Rarely	0	0%
Sometimes	45	90%
Often	5	10%
Total	50	100%

Table ten: TEFL tests repetition

There is a big problem in the TEFL tests may not only in TEFL test but tests in general which is repetition, in our sample 90% of them feel that the tests are the same of the last year and the year before and so on.

4. Do the TEFL tests provide any new cases to solve other than what have been discussed in the classroom?

Options	Subjects	percentage
Yes	30	60%
No	20	40%
Total	50	100%

Table eleven: the content of the TEFL tests

The results are a bit close, however; 60% think that test are giving some new cases to discuss; they have explained " the analytic part which is the essay is always about our opinion about some issues that happened to someone or about a new theory" (a student's explanation). However, the other 20% says that the entire test is about what we discussed in classroom. If not, the topics are closely related to what we can understand from the sessions.

5. The test should always include new topics that were not discussed in the classroom in order to evaluate the students' analytic ability?

Options	Subjects	percentage
Strongly agree	10	20%
Agree	20	40%
Disagree	2	4%
Strongly disagree	18	36%
Total	50	100%

Table twelve: the students' opinion about new topics in the test

It is obvious that even student can argue in the case of analytic test we can see that even if the agreed part of student is 60%, still a big deal of them disagree. The analytic part which holds most of time 40% of the test mark can be very difficult the student who focus on memorization where the other students might fail plus as a student explained the analytic questions take time to understand and answer.

6. "Should the methods of constructing tests be taught as separated module to insure that students can formulate effective test or it is sufficient to have it included under other modules such as didactics. In short passages please tell us your opinion."

The students' answers mostly agreed on this idea. They have explained that they do not need to learn about how to make test when they are teacher because that will affect at least the first student that they will teach. Taking a good classes about testing and practice how to make good tests for different type of students and classes can be the key to make better assessment in the future.

3.3. Analyzing the teacher's interview

To give credit and validity to our research we attempt after analyzing the tests and taking the opinion of the students we need to know how the teacher prepare and evaluate the tests that they made. We have directed two deferent types of interview to TEFL teachers (tow teachers) according to types of tests that re made in the English department at Biskra University.

1. "All the didactic tests that you constructed are much appreciated because they are well designed (the degree of difficulty" from easy to hard, from things students know to new subjects to discuss). Could you please tell us how and on what basis you construct your tests? "

"   All the didactic tests that you constructed are much appreciated because they are well designed. The first thing we noticed when analyzing your test is that you emphasis mostly on writing. Could you please explain to us the reason behind that?"

The first question in the interview was about how the teachers construct their tests. From their answer we concluded that most teacher attempt to use their own experience to mediate between what tests should contain and what the students can deal with. "Through years, the tests that I made have been similar to methodologists call piloting" (a teacher's answer) it is clear that through time teachers learnt to add, omit or modify some details in their tests. "We try to apply all the requirements of test but sometimes they do not work out well, simply because our students and administrations are not similar to those who proposed these requirements" (a teacher's answer). This answer show clearly that teachers encounter some challenges in preparing the tests that is related to the either the environment, the ways of learning or administration.

2.   "While analyzing the TEFL tests that were made by you, we noticed that you always include all the forms of test items in every test (multiple choice, true/false…). But for the student, there is always the issue of time. Could you please tell us how do you manage to keep the test in the proper length that student can understand and answer all the questions?"

"For every student writing a good paragraph is always a challenge especially when it is controlled by limited time. Could you please tell us how could you manage to construct tests that can fit the time that is given to the test?"

Time is always an issue to students; they can never have enough time to answer. The teachers explained how they have managed to make the test as long to make sure students will not have time to "cheat" in the main time the test can be answered in the

given time if the student is well prepared. An issue was raised by one teacher when she claimed that students misunderstand what is requested from them" the instruction is always clear (answer briefly) but for some reason they cannot have enough writing". This issue is real and can be noticed when revising the students' answer. But it is well known that when it comes to writing students always attempt to write as much as they can for the sake of covering anything that might crosses the teachers mind." We ask students to write to improve their writing skill but what we take in consideration is the content" (a teachers' answer).

3.   **"When it comes to the analytic part, the last part of your tests is the most negotiable subject by the student after the test. The analytic part is a perfect use of krashen's concept in learning (i+1), the student always needs to use their knowledge and some of what they can predict. Could you please tell us how could you construct the didactic tests using all the new testing method and in the mean time you keep your tests scorable for all students?"**

"In the analytic part student are always asked to be creative to some extent" (a teacher's answer). According to the teachers' answers it is curtail to push the student a little bit forward not for the sake of challenge them but to see whether they can take their understanding of the lessons to another level. It is not possible to through a whole new subject in the students' face during a very stressful time such as test but teachers attempt to use subjects that is partially related to what student have studied for example the lessons can be about listening this topic must include some uses of speaking even though these uses are just hints, the teacher can make a contribution about speaking in the test as an analytic topic.

4.   **"The new tests methods are emphasizing on that the test should carry up something new for the learners to deal with and to learn from it. Could you please tell us your opinion on this idea?"**

From the answers of teachers we can understand that it is possible to use some new issues but the teacher was a little bit cautious about using real life issues because the score for the answers will be more subjective and it is not good for the reliability of the test.

5.   **"Should the methods of constructing tests be taught as separated module to insure that students can formulate effective test or it is sufficient to have it included under other modules such as didactics. In short passages please tell us your opinion."**

The testing is a very interesting and important element is the evaluation of student but teaching it as a separated module is still early maybe for a start we can teach it as a whole chapter or ever a semester by including a lot of practicing but it does not have sufficient items to stand as a module.

When it comes to the teachers' opinion about test it is surly not the same as the student because from what we gathered teachers also encounter great difficulties that they try to work the best out of it. We also noticed that the experience plays a very important role in the teaching career staring from preparing test to the best way to evaluate student.

3.4.   General discussion

As we finished analyzing all the data gathered form the tools that we used, we will discuss their interrelationship between them in addition to the points of agreement or disagreement between the students' opinions and the teachers' clarifications if there are.

First, in terms of form; there is a general agreement that the form plays a great role in the test. It makes the test interesting which helps the test-taker focus on the questions. As it is important to prepare to the test it is also important to keep one's thoughts together during the tests, this problem may take place through many ways one of them the appearance of the paper which had a great appreciation among students and even among teachers. The order of the questions can help the students to keep their thoughts organized and focused on one topic and this can help students to a better use of time. Understanding the question is half of the answer; here where the instructions come to be highly needed in order to make the questions clear and easier to understand. Avoiding the students misunderstanding of the question is very crucial to have a better assessment to the test and this takes the reliability of test to a higher level. From the results that we have gathered, we also noticed that the use of different test forms can be sometimes a necessity; students usually have different appreciations to the forms where they can do better, also the teachers can cover a larger amount of the program using these forms.

Second, in terms of content; we have noticed a small gap between students and teachers; the students or most of them have strongly shown that they prefer to be tested only in what they have been taught. This can be due to the length of the program or the huge number of the students which make it hard to work with them all or evaluate them all. The teachers, on the other hand, want to evaluate all the aspects of the learning process starting from the progress of the program to what the students have understood and how they can use it. Both sides can be viewed to be right regarding to the good reasons that they gave their contributions. Unfortunately, none of them can change the fact that they are working in a posed environment where the teacher must finish the program and he/she needs to make sure that the program is well understood as well. The students in the other hand have to learn and succeed no matter what the difficulties he/she encounters.

To sum up,   after we have finished our research and our analysis we have noticed that both students and teachers have a good understanding of the situation that they are both in. This means that we have already made the first step to change which is knowing our weaknesses. Even though some of these weaknesses are posed, there is still good chance to find ways that shrink the gaps between what need to be done and what we are doing in the evaluating process.

Conclusion

In this chapter, we have gathered different types of data to cover all the angles possible about this the didactic testing in order to gain objectivity, validity and reliability. No difficulties were encountered during these operation thanks the cooperation of both students and teachers who showed a great maturity while dealing with our topic. As a conclusion, we found out that the testing topic had a great attention by the teachers who had been trying to transmit this attention to the students to create a new generation of teachers that may influence the future of testing in the English department of Biskra University. As researchers, we joyfully hope that this research may be a part of the change that the teacher and students seek for.

General conclusion

We have come from far away starting from defining term to gathering data about these terms trying to identify rules and basses that make good tests. Our work was directed to verify the strategies that are used in English branch of Biskra University; we also attempt to gather the teachers and students opinions about this crucial process in learning. For this sake, we managed to formulate some raised questions about the didactic tests in English branch of Biskra University. These questions were highly pointing on the didactic testing methods and their content. The aims of this study was to identify the weaknesses in

the methods used, what they lack then come up with some recommendations if possible. After all the discussion about the former, we attempt to be more precise in our study through developing head start hypotheses that we tried to decide whether they are correct or not. These hypotheses have limited our focus from the methods to the content. In our course of trying to answer our questions and verify our hypotheses we devoted two parts in our research a theoretical and practical part.

First, we have started with the theoretical part in which we have explained the differences between the common terms in language assessment that was used interchangeably by some linguists. Before start talking about the testing we was in need to bring a short history that shows the development of testing methods to put the reader in a clear image of what our subject is about. Validity, reliability, timing, test forms and ways of test were the spot lights which we tried to explain and use them as well in our research. The theory always became to put rules so we can work in a predicted way and arrive to a predicted result that is why we have used our theoretical part as a mediator and a means to give credit to our practical part.

Second, in our practical part we have given a high priority to document analysis as a start; in this matter we have tried to comber between our studies about testing methods and the tests that are made in English Department of Biskra Universities. The feedback of this analysis has provided a great deal of data that we have use in checking our hypotheses. The document analysis cannot stand alone as an argument so we have decided to direct a questionnaire for students and we made an interview with TEFL teachers. By the end of our research we had covered three angles: the test documents, the students' opinions and the teachers' arguments in addition of course the claims of other researchers that were discussed in the theoretical part. As researcher we have arrived to appoint that we can

check our hypotheses and decide whether they stand or fall when we expose them to our findings.

Before discuss the hypotheses we need to check the aims and the question of the research.(1) Evaluate the testing methods that are used.(2) Investigate whether the basic principles and rules of assessment are obeyed or not. (3) Identify the main advantages and disadvantages of the tests in didactics module. As former said, we had no attentions to evaluate the teachers because our concerns was only with the tests. Three goals were set and all of them were taken as a priority in the process of the research; we evaluated the tests and carefully compared them with our finding in the theoretical work then we identified some strength and weaknesses which were discussed former. We also set some questions for the sake of keeping the research in the right direction these questions werethe first to start with.

(1).Do the contents of the test serve the aims of didactics?. (2) What are the disadvantages of the testing methods that are used in didactics? (3) How can the appropriate contents in a test improve the student's awareness to the real problems of teaching-learning process? For the sake of answering these questions we divided our document analysis into two parts in sense of form and content. We also used some questions that were directed to the students and the teachers in order to give credit and objectivity to the research.

Finally, we come now to discuss our hypotheses which were (1).We hypothesize that the exercises used in the didactic tests are only focusing on recalling what has been discussed in the classroom.(2)We assume that didactic tests do not obey the basic principle of testing. These two hypotheses combined with the aims and the questions allowed us coming to some conclusions and they also raised some issues. We had found through our work that teachers DO provide new issues in their tests and further more they have a great

concern that these new issue must not be graded subjectively. The effort that was shown by our sample was enormous to prove that critical thinking is very important and also it should not be evaluated poorly. The most important thing that we have admired is when we asked about krushen learning strategy (i+1) the teacher had a great deal of knowledge not just about the topic but also how to use it. However; we also have noticed big deal of subjectivity in evaluating for some teachers, their answers ensured as that the tests was made to be for both the skill and the didactic knowledge. This method was encountering a very heavy criticism because it is simply unreliable. The issue that our research has raised is "the repetitive tests"; even for us as researchers this was a big problem because the document that we have analyzed was the same for three years. We can claim these tests as indirect cheating because the students become almost sure that the test will be the same as the last one, the test questions must stay unknown till the tests time come, but unfortunately this method does not serve this quest. Furthermore, this deed can cost the test its reliability and even its meaning because the results are already false.

To sum up, the strengths that we have found has provided us a great feeling of relief even if some of our hypotheses (We hypothesize that the exercises used in the didactic tests are only focusing on recalling what has been discussed in the classroom) were proved wrong. However, we had the chance to see the efforts that teachers made to give a good assessment and a good tests the only thing that we can prove in our hypotheses was the subjectivity that hold a big part in some tests. on the other hand, the test-takers or the students, we have noticed a great understanding about how tests should be made and how they should be evaluated; we can assume from this that teachers are doing a good job in the process of teaching which will make good teachers and good test makers in the future.

Recommendations

The research that we have started with general definition then we went deeper in exploring what other researchers have found around the world allow us to name the process of testing as "The art of testing". The art of testing does not rely only on theoretical rules but it also relies on the teachers experience and the identity of the test-takers.

After we have analyzed our research tools (documents, questionnaires and interviews) we have come to identify some problems and try to give some ideas that might help dealing with these problems. The first thing that we have noticed and we believe that it has a positive impact on the test reliability is the problem is the repetition of the tests; there are many ways now to prepare exercise for the test one of them is internet. The teacher might encounter some difficulties creating new tasks with time because of the curriculum the use of other tools can help the teacher and the students as well. We also advice the use of all test items, this can provide more flexibility in designing the test. The teacher always tries making sure to cover all the program in one test, with time this way also leads to repetition. We advise focusing on some point of the program and the rest of the lesson can be tested through other assessment tools such as homework; this way gives better assessment and more possibilities in making the test. The second problem was the length of the test; the essay item is known to be time consuming for the sake of better feedback. We highly recommend not to use it as one item test. If this item is important for some reasons such as promoting the writing style we recommend that this item do not exceed 50% of the mark otherwise the essay will affect the objective judgment of the test. The last point is, in our study we have encountered a problem that cannot be over looked which is the huge number of students Clay (2001) and Batchman (1995) explained that it is

impossible to use the essay test item in large number classrooms. The test in this case is simply unreliable; the best way to evaluate large classrooms is multiple choice test items.

To sum up, our research was taken to verify some hypotheses some of them were rejected (the teacher do not include new topics that helps students to deal with real teaching problems) and others were not (the teacher do not obey the basic rules of didactic tests)this hypothesis was partially proved correct because some teachers do respect these rules . Through this study we have found some new problems which we did not attempt to deal with however the research first principle is to find problem, we wish that these problems will be solved in the future.

Students' Questionnaire

Dear students, you are kindly requested to fill in this questionnaire which will take approximately 25-30 minutes to complete. We are carrying out an investigation about

Strategies to Construct Useful Tests. Your answers are very important for the validity of the research we are undertaking. As such, we hope that you will give us your full attention and interest. Please, mark (√) the appropriate box (es) or give full answer(s) where the gaps are provided. 50

Thank you, in advance, for your collaboration.

Section One: General Information

1.   How can you describe your level in English language?

a.   Average

b.   Good

c.   Very good

d.   Excellent

2.   On what basis do you build your self-evaluation?

a.   Tests' marks

b.   Language understanding

c.   language use

d.   Others

If others, ……………………………………………………………………

3.   What do tests mean to you?

a.   Marks and success   ☐

b.   Identifying weaknesses and strengths   ☐

c.   Contest to compete with your class mates   ☐

d.   A chance to learn and improve your abilities ☐

e.   Others   ☐

If others, …………………………………………………………………………..

……………………………………………………………………………………………

4.   Do tests reflect the real level of students?

a.   Yes   ☐   b. no   ☐

 Please explain

……………………………………………………………………………………………

……………………………………………………………………………………………

5.   Which of the following test forms you prefer more? (You can tick more than

one)

a.   Multiple choice   ☐

b.   True/false   ☐

c.   Matching   ☐

d.   Fill-in-blanks   ☐

e.   Essay   ☐

Why,…………………………………………………………………………………………

…………………………………………………………………….....................................

....

6.   Do you appreciate the test to be mixture of all forms?

a.   Yes   ☐   b. No   ☐

Explain, …………………………………………………………………………………

……………………………………………………………………………………………..

7.   Do you think that tests are the best way to evaluate and decide whether student succeeded or failed?

a.   Yes   ☐   b. No   ☐

8.   Please, explain why

……………………………………………………………………………………………

……………………………………………………………………………………………

………………………………………………………………………………………

Section Two: TEFL and tests

1.   Do you have difficulties in TEFL tests?

a.   Yes   ☐   b. No   ☐

2.   If yes what kind of difficulties you often encounter ?

a.   Time   ☐

b.   The content   ☐

c.   The language used ☐

d.   Others   ☐

If others, Explain why

………………………………………………………………………………

……………………………………………………………………………….

3.   How often do you feel that TEFL tests are repetitive?

a.   Never   ☐

b.   Rarely   ☐

c.   Sometime   ☐

d.   Often   ☐

4.   Do the TEFL tests provide any new cases to solve other than what have been discussed in the classroom?

  a. Yes   ☐   b. No   ☐

Explain, ………………………………………………………………………

………………………………………………………………………………………

5. **The test should always include new topics that were not discussed in the classroom in order to evaluate the students' analytic ability?**

a. Strongly disagree ☐   b. Disagree ☐

c. Agree ☐   d. Strongly agree ☐

Why,

………

....

6. **Should the methods of constructing test be taught as separated module to insure that students can formulate effective test or it is sufficient to have it included under other modules such us didactics. In short passages please tell us your opinion.**

………………………………………………………………………………………………………

………………………………………………………………………………………………………

………………………………………………………………………………………………………

………………………………………………………………………………………………………

………………………………………………………………………………………………………

………………………………………………………………………………………………………

………………………………………………………………………………………………………

………………………………………………………………………………………………………

………………………………………………………………………………………………………

"Thank you for your cooperation "

Teachers' interview

Dear teacher (1), you are kindly requested to fill in this open-ended questionnaire. We are carrying out an investigation about Strategies to Construct Useful Tests. Your answers are very important for the validity of the research we are undertaking. As such, we hope that you will give us your full attention and interest.

Thank you, in advance, for your collaboration.

The questions:

1.   All the didactic tests that you constructed are much appreciated because they are well designed (the degree of difficulty" from easy to hard, from things students know to new subjects to discuss). Could you please tell us how and on what basis you construct your tests?

……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

………………………………………………………

2.   **While analyzing the TEFL tests that were made by you, we noticed that you always include all the forms of test items in every test (multiple choice, true/false…). But for the student, there is always the issue of time. Could you please tell us how do you manage to keep the test in the proper length that student can understand and answer all the questions?**

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

……………………………………………………………………………………………

3.   When it comes to the analytic part, the last part of your tests is the most negotiable subject by the student after the test. The analytic part is a perfect use of krashen's concept in learning (i+1), the student always needs to use their knowledge and some of what they can predict. Could you please tell us how could you construct the didactic tests using all the new testing method and in the mean time you keep your tests scorable for all students?

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

………

………

………

………

………

………

………

………

………

………

………………………………………………………………………………

4. **The new tests methods are emphasizing on that the test should carry up something new for the learners to deal with and to learn from it. Could you please tell us your opinion on this idea?**

………

………

………

………

………

………

………

………

………

………

………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

……………………………………………………………………………....

5. Should the methods of constructing tests be taught as separated module to insure that students can formulate effective test or it is sufficient to have it included under other modules such as didactics. In short passages please tell us your opinion.

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

Note: We kindly request to aid us with some copies of your old didactic tests.

"Thank you for your cooperation "

Teachers' interview

Dear teacher (2) you are kindly requested to fill in this open-ended questionnaire. We are carrying out an investigation about Strategies to Construct Useful Tests. Your answers are very important for the validity of the research we are undertaking. As such, we hope that you will give us your full attention and interest.

Thank you, in advance, for your collaboration.

The questions:

1.   All the didactic tests that you constructed are much appreciated because they are well designed. The first thing we noticed when analyzing your test is that you emphasis mostly on writing. Could you please explain to us the reason behind that?

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

……

……

……

……

……

……

……………………………………………………………

2.   For every student writing a good paragraph is always a challenge especially when it is controlled by limited time. Could you please tell us how could you manage to construct tests that can fit the time that is given to the test?

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……………………………………………………………………………………………

3.   When it comes to the analytic part, the last part of your tests is the most negotiable subject by the student after the test. The analytic part is a perfect use of krashen's concept in learning (i+1). The student always needs to use their knowledge and some of what they can predict . Could you please tell us how could you construct the didactic tests using all the new testing method and in the mean time you keep your tests scorable for all students?

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

………………………………………………………

4.   The new tests methods are emphasizing on that the test should carry up something new for the learners to deal with and to learn from it. Could you please tell us your opinion on this idea?

……

……

……

……

……

……

……

……

……

………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………

5. Should the methods of constructing tests be taught as separated module to insure that students can formulate effective test or it is sufficient to have it included under other modules such as didactics. In short passages please tell us your opinion.

………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………

……

……

……

……

……

……

……

……

……

Note: We kindly request to aid us with some copies of your old didactic tests.

"Thank you for your cooperation "

Mohammed Kheidher University of Biskra

Faculty of Arabic Language Arts and Literature & Foreign Languages

Division of Foreign Languages

Department of English Studies

(Option) Post-Graduation   (Major) Sciences of Language

(Module) Language Teaching Methodology   (Time Allotted) 90 mns

Date : 1.21.2015   (Instructor) Dr. Bashar, A.

First Term Examination Paper in Language Teaching Methodology

NB : *(1) Section One is Compulsory*   *(2) Choose either Section Two or Section Three.*

Section One : Let's Find Out (9/9)

Task One

Are the following statements true (T) or false (F) ? Write T/F in the space provided. (2.5/2.5).

1. _T_ . Intensive reading is usually indepth, in-class reading.

2. _F_ . Schemata are predictable sequence of events.

3. _F_ . Critical listening refers to the type of listening where students listen to appreciate listening materials (videos, movies, songs, and interviews, etc.).

4. ___ . In reading, accuracy and fluency are often used interchangeably.

Task Two

Peruse (i.e., read) the definitions on the right and insert the right term in the space provided (2.5/2.5)

1. _Extensive_ It is a kind of out-of-class reading for pleasure usually implying a good deal of reading.

2. ___ . It is about doing tasks that require authentic communicative language use (Such tasks can include visiting a doctor, conducting an interview, or calling customer service for help.

3. _____ . It is the ability to hear, identify, and manipulate individual sounds-phonemes--in spoken words.

4. _Phonics_ . It is the study of the relationships between letters and sounds that they represent.

Task Three

Briefly Explain the following statement (1/1).

I really appreciate your critical discourse.

Task Three (Cloze Test)

Read the passage below and fill in the gaps with the words in the box (5/5)

Protocol- reading- strategies- reading -comprehension- reading- reciprocal- readers- feedback- reading

If, as instructors of __(a)__ , we want to teach our students to be active __(b)__ , __(c)__ teaching (Palinscar and Brown 1984) gives students the required cognitive and metacognitive __(d)__ to understand the __(f)__ process. __(g)__ teaching provides models and requires interaction and __(h)__ to engage learners and reinforce comprehension in jointly reconstructing the meaning of text. Although Palinscar and Brown's original study involved seventh grade English speakers who were recognized as having poor __(i)__ skills, the reciprocal teaching __(j)__ is applicable wherever readers are vulnerable to comprehension failure.

Palinscar and Brown selected four skill areas to activate and monitor __(k)__ comprehension: *(a) summarizing, (b)questioning, (c) clarifying, and (d) predicting.* They state:

By asking students to summarize a section of text, the teacher is requesting them to give attention to content. They are then asked to compose questions about that content in order to clarify meaning and engage in critical evaluation. Finally, the teacher asks the students to make predictions to involve them in drawing inferences. All four of these steps activate relevant background knowledge.

Section Two : Let's Analyze (10/10)

Read the passage below, then answer the questions all the more backing up from the text.

The controversy about the role of grammar in L1 writing instruction in some sense forced researchers and practitioners to reconsider the relationship between grammar instruction and language accuracy in student writing. This is most evident and helpful in the publication of Mina Shaughnessy's groundbreaking work, *Errors and Experimentation* (1971), in which she redefined the notion of error. In her research and teaching, she found that native-speaking student writers reading their compositions aloud could actually correct (most) errors without realizing they had made them. She concluded that contrary to teachers' assumptions, students don't make errors because of an incomplete grasp of the rules of English grammar ; in fact, they have an intuitive grasp of many grammar conventions. When they do make errors, these are often performance errors or ones based on "rules" about written language that they have erroneously intuited. Shaughnessy's work and that of subsequent researchers in the field of basic writing have encouraged writing teachers, who may have been prone to prescriptivism, to view error as a window on students' development as writers.

Barbara Kroll (2003). *Exploring the Dynamics of Second Language Writing.*

Global Comprehension (4/4)

1. What is Barbara Kroll's point in this passage ?

2. What does the field of EFL/ ESL writing owe Mina Shaughnessy ?

3. What is *performance error* ?

4. What is Mina Shaughnessy's groundbreaking finding ?

Detailed Comprehension (8/8)

Do please note that the cluster of the following questions requires of you to read between the lines to be able to fully and insightfully answer them.

1. Why do researchers and practitioners have to reconsider grammar-writing relation ?

2. What is meant by "intuitive grasp of grammar conventions" ?

3. What is the difference between native and non-native speakers' writing errors ?

4. In what way are prescriptivist and descriptivist teachers similar and/ or different ?

Section Three : Let's Think Critically (10/10)

Topic

The assessment of L2 writing, which seems to be a problematic task for most teachers, is carried out on two different levels : Form- and Content-Level. On the form-Level, the mechanics of writing, grammar, spelling, etc. are the focus of teacher's assessment. On the content-level, the emphasis is on organization, argumentation, ideas, etc.

Instructions

Rely on your own experience as a Master's II student and would-be teacher, answer the following questions : Which one of either assessment is more beneficial for Algerian EFL learners ? Why ? What are your suggestions to make L2 writing assessment more likely to be objective, reliable, and comprehensive?

End of the Examination Paper

References

Bachman, L. F. (1995). *Fundamental Consideration in language testing*. Oxford: OUP.

Bachman, L. F. (1999). *Language Testing at the Turn of the Century: Making Sure that What We Count Counts*. AAA Letter, 20(2). Retrieved from the World Wide Web:http://aaal.lang.uiuc.edu/ letter/testing.html

Boyle, J. &Fisher, S. (2007). *Educational Testing: A Competence-Based Approach.* Oxford: BPS Blackwell.

Brown, J. D. (1996). Testing in language programs. Upper Saddle River, NJ: Prentice Hall.

Clay, B. (October,2001). Is This A Trick Questions. Kansas Curriculum Center.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. Applied Linguistics,

Carroll, J. B. (1972). Fundamental considerations in testing for English language proficiency of foreign students. In H. B. Allen, & R. N. Campbell (Eds.), Teaching English as a second language: A book of readings (2nd Ed.). New York: McGraw-Hill.

Council of Europ.Manual for Language Test Development and Examining.

Cumming, C. W.-S. (2009). Educational Assessment in the 21st Century. Springer.

David I. Hanauer, G. F.-S. (2009). Active Assessment :Assessing scientific Inquiry. New York: Springer

Davidson, G. F. (2007). Language Testing And Assessment. London: Routledge. Deygers,

Davis, Barbara G. "Quizzes, Tests, and Exams," Tools for Teachers, Jossey-Bass (1999)

http://uga.berkeley.edu/sled/bdg/quizzes.html

Deygers, B. An introduction into (English) Language testing.Ghent University.

Dunlap, C. Z. &Weisman, E. M. (2006).Helping English Language Learners Succeed. Huntington Beach: Shell Education.

Fulcher, G. (2010). Practical Language Testing. London: Hodder Education.

Fulcher, G. & Davidson, F. (2007).Language Testing and Assessment Abingdon: Routledge.

Geoffrey Broughton, C. B. (2003). Teaching English as a Foreign Language. London: Routledge.

Germains, P. R.-D. (1992). Evaluation. Walton Street-Oxford OX2 6DP: Oxford University Press.

Gupta, K. (2007). A practical Guide to Needs Assessment, 2nd ed. San Francisco:

Pfeiffer.

Heaton, J. B. (1988). Writing English Language Tests. Longman Group UK, Limited.

Hinofotis, F. B. (1981). Perspectives on language testing: past, present and future. Nagoya Cakuin, Kyoikukiyo. . http://www2.hawaii.edu/~roever/wbt.htm

Johnson, John A. "Levels of Understanding Assessed by Multiple Choice Questions"

   http://www.personal.psu.edu/faculty/j/5/j5j/psy002/MCQ.html

Martha L. A. Stassen, K. D. Program -Based Review and assessment. UMASS.

Nunan, D. (2004). Task-Based Language Teaching Cambridge: Cambridge University Press.

Oller, J. W., Jr. (1979). Language tests at school: A pragmatic approach. London: Longman.

Ozerova, A. (2004). Types of Tests Used in English Language.University of Latvia press.

Rea-Dickens, P. & Germaine, K. (1993).Evaluation. Oxford: Oxford University Press.

TUNC, F. (2010). Evaluation Of An English Language Teaching Program At A Public University Using CippModela. Middle East Technical University.

Westwood, P. (2005). Spelling: Approaches to Teaching and Assessment, 2nd ed. Camberwell: ACER Press.

Wyatt-Smith, C. &Cumming, J.J. (2009).Educational Assessment in the 21st Century: Connecting Theory and Practice. New York: Springer.

ملخص

لقد قمنا بهذا البحث من اجل التاكد من ان اساتذةالديداكتيك في فرع اللغه الانجليزية بجامعه بسكرة يحترمون قواعد تصميم الامتحانات ام انهم يعتمدون على خبرتهم فقط. تم هذا البحث سنة 2015 من اجل تحليل عينات من الامتحانات التي تم تصميمها من قبل اساتذة الديداكتيك في مختلف السنوات. كما قمنا بجمع المعلومات الخاصة براي الطلبة عن طريق تقديم اسئلة لمجموعه من طلبة الماستر 1 انجليزية . و بغرض اعطاء هذا البحث مصداقية اكبر وجهنا بعض الاسئلة لاساتذة الديداكتيك لاخذ فكرة عن كيفية قيامهم باعداد الامتحانات . تحليلنا للامتحانات اظهر ان الاساتذة يهتمون بقواعد اعداد الامتحانات لكن الحوار الذي قمنا به معهم اظهر انهم يعتمدون بشكل اكبر على خبرتهم الشخصية , من جهة اخرى اظهر الطلبة ان هاته الامتحانات مناسبة لهم بشكل او بآخر . في هذا البحث قمنا بتحليل عينتين فقط من الامتحانات و هذا راجع الى ان الامتحانات في لثلاث سنوات. الماضية كانت متشابهة.