

WRITTEN TESTS APPLIED IN THE ENGLISH AND FRENCH AND ENGLISH
AND SPANISH PROGRAMS AND THEIR RELATIONSHIP WITH THE
PRINCIPLES OF LANGUAGE ASSESSMENT

By

Mario Gustavo Rivera Ortega
Luis Eduardo Salas Riascos

Human Science School
Linguistics and Languages Department
English and French Program
University of Nariño
November, 2009

WRITTEN TESTS APPLIED IN THE ENGLISH AND FRENCH AND ENGLISH
AND SPANISH PROGRAMS AND THEIR RELATIONSHIP WITH THE
PRINCIPLES OF LANGUAGE ASSESSMENT

By

Mario Gustavo Rivera Ortega
Luis Eduardo Salas Riascos

ADVISOR

Mg. Ana Clara Sánchez Solarte

Human Science School
Linguistics and Languages Department
English and French Program
University of Nariño
November, 2009

Qualification

Signature

San Juan de Pasto, November 17, 2009

ACKNOWLEDGEMENTS

Special thanks to professors Gabriel Obando and Vicente Araujo for their knowledge and help.

Thanks to our family and their support.

And especially to professor Ana Clara Sanchez for her knowledge and dedication.

DEDICATED TO

*I am enormously pleased for
Sharing this success
With my family,
And all the people I love.
Luis Eduardo.*

DEDICATED TO

*In loving memory of my father,
I am grateful for my mother
Thank you for your support and love.
Thanks also to my teachers for their patience
Dedicated to my beloved son*

Mario R.

TABLE OF CONTENTS

ABSTRACT

RESUMEN

RESEARCH PROBLEM

Research Problem	12
Problem Description	12
Problem Statement	14
Purpose and Objectives	15
Purpose of the study	15
Specific Objectives	15
Significance of the Study	15
Delimitations and Limitations	17
Delimitation	17
Population Delimitation	17
Geographical delimitation	18
Temporal Delimitation	18
Conceptual Delimitation	18
Limitation	19

LITERATURE REVIEW

Learning Theories	21
Behaviourism	21
Cognitive Theory	23
Krashen (monitor model)	24
Acquisition Learning Hypothesis	24
Monitor Hypothesis	25
Natural Order Hypothesis	26

Input Hypothesis	27
Affective Filter Hypothesis	28
Innatism	29
EVALUATION	31
Testing and Assessment	31
Proficiency	34
Reasons for Classroom Testing	42
Test design	43
THE RESEARCH METHOD	
The Design	45
Types of research	46
Subjects or Participants, Setting And Sample	46
Subjects	46
Setting	47
Sample	48
Procedure	48
Variables or Factors	48
Data Collecting Techniques and Materials	49
Description of the Procedure	49
Validation Criteria	50
Pilot Study	51
Data Analysis and interpretation	51
Ethical Issues	52
Analysis and interpretation of data	54
Conclusions and recommendations	81
Conclusions	81

BIBLIOGRAPHY

APPENDIX

ABSTRACT

If we are going to teach other people a new language, we should also know how that teaching and learning process is going to be assessed.

This study was focused on a subset of assessment which is testing, and how this process is carried out at the University of Nariño, especially at the English and French, and English and Spanish programs.

The purpose of this study was to find out if the principles of language assessment are being applied in written tests taken by learners. The instrument we used to achieve this aim was a survey applied to teachers and another to students in order to compare the perception they have about the written tests already taken.

The results provide information about the principles teachers apply when assessing learners and some suggestions and comentaries about the principles which must be used in a more appriate way.

RESUMEN

En el proceso de enseñanza de una segunda lengua es importante tener en cuenta la forma como se llevara a cabo la evaluación de dicho proceso.

Este estudio se enfatizó en un componente de la evaluación el cual se refiere a los instrumentos utilizados para comprobar el manejo de una segunda lengua, en especial como se lleva este proceso en los programas de Inglés-Francés y de Inglés-Español de la Universidad de Nariño.

El propósito de este estudio fue ahondar si los exámenes tomados por los estudiantes siguen los principios de evaluación propuestos por los autores citados en este proyecto. Para lograr este propósito se desarrollaron dos encuestas, una dirigida a estudiantes y otra a profesores con el fin de comparar la forma con la cual observan el desarrollo de este proceso.

Los resultados nos suministran información acerca de la forma como los profesores incluyen los principios al evaluar a sus estudiantes. De la misma manera sugerimos aspectos importantes en la creación y aplicación de dichos exámenes.

Research Problem

This chapter will be about the main objectives and purposes of the project, and how this project will be useful for our society, and especially for the Linguistics and Languages Department where this paper is going to be carried out. In order to achieve these purposes we can say that it is important to delimitate this proposal for obtaining trusting results at the end of it. Also, it has been taken into account that this proposal has some limitations for its development.

Problem Description

Language learning and teaching are central topics in the academic community; in every place where English is taught, there must be a process of assessment, and it is important to analyze how this process is being carried out; considering that the knowledge of a second language is a necessary element in order to comply with the high requirements in the society.

According to a personal observation, it is found that some written tests do not measure what they are intended to, and as a result of this, it indicates low performance after being applied. This experience is a starting point for this research, because it was considered that it is an important problem to be analyzed.

In order to give a contribution to the analysis of this problem, it has been decided to focus on the

instruction aspect, and especially in the testing process, which is immersed on it and which is a very interesting aspect that requires a deep study, especially a study based on the principles of language assessment -which will be the specific point of this investigation.

During many years, testing process, in teaching languages, has been a difficult tool in terms of management, because teachers usually do not know how they should prove or measure students' capacities. The most difficult aspect in testing has been the formulation and preparation of tests, so we think that it is very important the contents and the forms how tests are done; according to Backman, F. (as cited in Tim McNamara, 2001, p. 95) "one of the characteristics of a test that we as a prospective test users examine is its content. As we can see the content in the design of tests is very important, specially for students who should have a clear vision of what they are going to be evaluated on.

we can also say that testing measures and verifies how much knowledge students have achieved at the end of a unit or course, so testing becomes an important instrument for teachers which contribute in many ways to the evaluative process, not only for the educative development of students but also in the development of different aspects of their lives. McNamara, T. (2001 pg. 3)

"testing is a universal feature of social life:

Throughout history people have been put face tests to prove their capabilities or to establish their credentials''. So through tests teachers or assessors can see their students in an integral way, and tests have to be characterized according to their reliability and the validity.

We has considered that it is very important to know how tests are been formulated and also their characteristics in order to discover if written tests are making a contribution to the language learning or if they are a problem for students in the English and French and English and Spanish programs at the University of Nariño.

As we can see the performance in the English language can present some problems which require of a deep study, besides we can realize that this situation can be influenced by many factors specially one of them which will be the micro problem of this investigation. Thus we want to research how written test are being designed and if they are based on the principles of language assessment and testing, or the approach they are based on.

In this way, in the next part of this writing we are going to refer to these factors in a more specific way.

Problem Statement

Do the written English tests applied in both the English and French and English and Spanish program from the

Human and Science School at the University of Nariño;
follow the principles of language assessment?

PURPOSE AND OBJECTIVES

Purpose of the study

To analyze if written tests are being designed with the main principles of language assessment, and probably how they can be changed in order to obtain more appropriate information about student's knowledge or domain of a second language.

Specific Objectives

To examine if some of the principles of language assessment, cited in this paper, are being applied at setting written tests.

To analyze if the contents on the syllabus provided by teachers are being measured with the test applied.

To collect information about the frequency of time teachers renew or update written tests when they are in charge of the same subject more than once.

To identify the possible limitations on written tests.

Significance of the study

It is worth stating at this point that assessment has been a topic of research during a long period of time and we can consider necessary to investigate a subset of the evaluation process which is testing. Let us consider that tests are instruments of measuring personal abilities, in this way tests become in a part of the educative community

because they measure and show the degree of understanding of a topic.

The way how tests are designed and prepared can be directly related to the way they are going to be developed by students and if they do not have the main characteristics of a good test they will not show what the real results of the teaching process and students' progress and performance are.

Moreover, it is important to mention that tests must be analyzed before and after they have been applied, and determine if they are achieving their purpose or if they are only part of the syllabus or the program.

According to our experience as English students, some final or mid-terms tests applied at the English and French program are not measuring what they are intended to. For this reason we consider that they should be redesigned or that they must be analyzed in order to know why these problems can be observed.

Nowadays, the English and French program has a subject in Ninth semester which is named "Assessment". This subject did not exist in the previous curriculum, likely this is a factor which can be influencing in this problem because some teachers, who studied with the previous curriculum, might be lacking the knowledge in some of the requirements of a written test.

On the other hand, we can consider that this proposal could be doable because we count on the willingness and collaboration of teachers, students and administrators of the Language Center where one will develop this project, in the same way, we can dispose of elements and resources to achieve our purposes; we also disposes of time in order to analyze and collect all the information as much as possible in order to obtain clear results.

Finally, we can say that this investigation is feasible due to many researches carried out about it that say that testing is important. For example, according to McNamara(2001)'' language tests play an important role in many people's lives, acting as gate ways, at important transitional moments in education...''. As we can see an investigation about testing could be useful in order to contribute to many fields of education and society.

DELIMITATIONS AND LIMITATIONS

Delimitation

Population delimitation

For this investigation the ninth semester of the English and French and English and Spanish program from the human Science School at the University of Nariño will be considered. As well as the teachers in charge of these courses.

Geographical delimitation

This investigation will take place in the Language Department at the University of Nariño Panamericana campus, which is located in the north part in the city of Pasto.

Temporal delimitation

This research will be developed between September 2009 and November 2009. During this period of time we will collect the necessary data and analyze it, at the end of this time we want to have a clear result about what it has been found and its comments and suggestions.

Conceptual delimitation

According to H. Brown, (2002), a test is a method of measuring a person's ability, knowledge or performance in a given domain.

According to Brown (2002) a test is practical when it is not excessively expensive, besides it should stay with appropriate time constraints, it is relatively easy to administer and it has a scoring evaluation procedure that is specific and time-efficient.

Brown, 2002 affirms that a test is reliable when it is consistent and dependable; it means when applying the same test to different audiences it should provide similar results.

VALIDITY: it is the extent to which inferences made from assessment results are appropriate, meaningful, and

useful in terms of the purpose of the assessment Gronlound (as quoted in Brown, 2002)

WASHBACK: according to Hughes (as quoted by Brown, 2002) it is the effect of testing on teaching and learning, washback also refers to the effects the tests have on instruction in terms of how students prepare for the test.

AUTHENTICITY: Palmer and Bachman (as quoted by Brown, 2002) it is the degree of correspondence of the characteristics of a given language test task to the features at a target language task.

ASSESSMENT: According to Brown (2002) it is an ongoing process that encompasses a much wider domain. Whenever a student responds to a question, offers a comment, or tries out a new word or structure, the teacher subconsciously makes an assessment.

CONTENT: According to Hedge (2000) it is broader, with increased emphasis on the sub-skills. The content may be focused on a learner's ability to extract meaning.

FORMAT: Hedge (2000) affirms that formats have gone from explicitly language-focused tasks to tasks which are communicatively contextualized.

Making criteria: these are visible for spoken interaction and written language, Hedge (2000).

Limitation

We have considered as possible limitation to this project the lack of motivation by teachers for answering

the questions of the interviews, also we think that some external factors related with the University of Nariño as protests, disturbs and other ones would affect the normal sequence of the investigation.

In this chapter we have talked about the performance of the English language in the Language Center at the University of Nariño, in this way we give some reasons, aspects and characteristics about it, and it has been identified the aspects or variables that affects it, it also talked about the importance, the influence and the relation among these variables and the macro problem, especially we talked about the importance of the types and characteristics of tests which is our micro problem of investigation or independent variable. On the other hand, we specified our problem of investigation in terms of objectives, limitations and delimitations. In the next chapter we will specify aspects about the literature review of this project.

LITERATURE REVIEW

The human being has always been concerned with how the brain acts, but in this process some authors have dedicated their studies to a simple thing, how the human learns and teaches something, and especially how the human learns and teaches the language.

Language testing is a tool which provides teachers with the information about the students' domain of a certain topic in L2 after a teaching process, for this reason it is important to know how we learn in order to know how to evaluate. Here, one can find some theories which try to explain the learning process of a second language and also you will find a brief description of our project of investigation.

LEARNING THEORIES

Behaviourism

According to Lightbown and Spada, (1999) behaviorism may explain how the first contact between children and the language is produced. Also, this theory could allow us to know some of the most complex aspects of the language in children's mind which can influence on the process of language learning.

Behaviorism shows the learning process as the sequence of imitation, practice and habit formation. Children imitate sounds and patterns which they hear around them, and in this way, imitation is a sequence of acts which are

repeated and they are based on a model. Thus, these acts become a practice, it means that learners apply or repeat the model which has been presented, after this process "feedback" on success takes places, and in this way, learners have the opportunity to know the errors in order to correct them, however the behaviorists, in language acquisition, say that imitation and practice alone cannot explain the most common language productions made by children because they receive information and then they generalize that information in new contexts. Also they create new forms and new ways of presenting the data and finally they try to develop a sequence of habits that will allow them acquire the most complex structures of the language.

On the other hand, according to the behaviorists, this theory is connected to the Contrastive Analysis Hypothesis that mentions the learner as a person who will acquire the structures of language with facility if they are nearly alike to his first language, but the learner will find difficulties in the structures that are dissimilar.

As in first language acquisition the behaviorism has proven to be at best an incomplete explanation of second language acquisition. Psychologists have moved on to new more complex theories of learning.

Cognitive theory

According to Lightbown and Spada, (1999), the cognitive theory sees the acquisition of the second language as the improvement of the knowledge system where students make an effort to absorb the components of the language that they want to produce in the practice, this process is developed in a conscious way as a result of the interaction between the information that is available in the environment and the person who absorbs that information, so in the course of time the learner will have the capacity of using those components in a mechanical way. As we can see, the cognitive theory focuses on mental aspects in order to learn a second language, for instance, it takes into account the way learners perceive the knowledge, the capacity that learners have to memorize that knowledge, the learner's intelligence, etc. In order to explain these mental aspects the cognitive theory attends to multiple ways, and one of them is the development of the information, the way how they are caused, and how they are built by each one. However, cognitive psychologists (as cited by Lyghtdown and Spada, 1993), say that there is a visible fact called "restructuring" this fact states that the things that we know and which are improved through practice are not explainable, because they see acquisition of new knowledge as a system supported on the participation of the knowledge which without practice may be

"restructured", in this way students will find an important advance or a retrogress when the systematic aspect of learners is assimilated in incorrect way.

In our opinion, this theory shows how the learner makes use of the elements that are useful for him or her, and he or she can develop his or her abilities through them. In this way second language learning may be a mental product. This theory can contribute for ESL, because it takes into account the internal process of the learner, and it is expressed in different kinds of production with the use of the language. This theory can also help us for understanding what are the external elements that learners put into practice in the development of his or her abilities in the process of testing.

Krashen (monitor model)

According to Stephen Krashen (as cited in Lightbown & Spada, N., 1999), there are five specific hypotheses, which try to explain the learning process; these hypotheses constitute the monitor model, and they are:

Acquisition learning hypothesis

According to Krashen (1982) acquisition and learning are the two ways for students to approximate learning a second language, in this way learning means that a student knows the components of the language through the reasoning, the use of sources of information. On the other hand, acquisition is a process which is not intentional; it is

natural for each person who is unconscious about the interaction with the second language.

This hypothesis gives us clear concepts about what is learning and what is acquisition, and this is an important starting point for the comprehension of terms and processes, besides, we consider to know to recognize and determine what is the difference between these concepts. Also this theory is useful because many people think that learning and acquisition are the same aspects, and this theory helps people in order to obtain a clearer vision about both processes.

Monitor hypothesis

Krashen (1982) affirms the existence of an organized system that works as a natural monitor or editor. This system gets the information of the environment and it establishes a process which chooses and selects of a natural way the correct information that a learner should keep in his or her mind. On the other hand, this system focuses its use on content rather than form, it means that the structure of the words, sentences, etc, it does not matter at the moment of selecting the information because Krashen (1982) states that the knowledge of the rules helps the speaker to refine what he has acquired through communication in a real context, and for this reason language learning should be communicative and not only a grammar learning process.

We can say that this theory can contribute to the communicative process rather than in the grammar one, and it may be a more appropriate way to learn a second language. In spite of the fact that it is not possible to affirm the existence of a physical monitor in the mind, it is seen just as a mental process that helps people to develop a learning process; however this theory is important for us because we think that when this theory affirms the existence of a natural monitor or editor system it means that everybody has a program which selects and chooses the correct information from the environment in order to keep a real and a knowledge without errors and it can be maybe impossible, so we consider that this theory is very interesting in order to analyze it .

Natural order hypothesis

Krashen (1982) affirms that we acquire the rules of a language in sequences, which are acquired early, while others are acquired later. In addition the easiest rules or structures are not necessarily the first to be acquired and the natural order is independent of the order in which rules have been taught. It means that the human beings have a natural program which gets the rules and the structures of the language in specific order. Specific because this theory affirms that the rules or structures are acquired from the easiest to the most difficult rules, so the effort that this program makes is just on the form rather than the

content of the language. In this way, learners are going to obtain a big knowledge in the grammar part of that language and then they are going to put in practice that knowledge in a real context.

Usually, each process of learning has a sequence or natural order, and we consider that in the language learning process, there is always a specific order of learning, especially in grammar, so the idea of this theory can be useful for second language learners because according to it we acquire a language by steps, the first one is the acquisition of the grammar in an specific order, and the second is put in practice that grammar into communication. In this way we consider that this theory is also very interesting because it needs of a deep study and analysis.

Input hypothesis

In this view, Krashen (1982) affirms that the language which a learner understands may be comprehensible in part, because of clues such as gestures, situations or prior information, so we acquire a language by receiving comprehensible input, but it is not the enough condition for acquisition. It means that the environment which contains the information that we are going to acquire is not the isolated element that a learner needs in order to acquire a language, because he or she needs of a series of rules with the other components that complete the learning

process, so in this way, the input is an essential element but, it is necessary that the input contains clear information. That means that the input is comprehensible to the learner, because it can help students with correct information, and in a correct way, to the contrary it can be confusing, and in this way it does not contribute to the acquisition of a second language.

This theory can contribute to the teaching process, because it shows us that it is important to take into account the most comprehension way of teaching, and in this way intake and output will be successful.

Affective filter hypothesis

Here Krashen (1982) focuses his study on the internal or emotional part of the learners which is something very important, he affirms that the conditions or the states of the learner affect the acquisition of a second language because the learner is tense, anxious or bored, and he or she will screen out the input making it a barrier for acquisition, this hypothesis has implications for classroom practice, because teachers cannot understand in a better way their students' performance and also recognize the reasons which some students have different levels, and the reason can be that they present different emotional states that can affect them at the moment of the acquisition, in this way the feelings have a direct incidence on the knowledge because the emotional states can create a

vulnerable situation in the mind of the learners and in this way it can be a big obstacle for the acquisition.

We consider that this theory explains some important causes in relation to academic performance, in our case why learners feel fear about testing process; in this way one must take into account this theory as a guide for our proposal because the way in which testing is being produced can become a real problem for producing knowledge in the test takers. It also can help us to understand how the teachers' feelings can create situations in which the learner can develop some emotional states.

Innatism

This theory makes the assumption that everybody has already fundamental information about languages and because of this it is possible to learn the specific language of the environment. According to Chomsky (as quoted in Lyghtdown & Spada, 1994), the language of a child who is in contact with the environment is full of disorganization and confusion, and parents can be an element which helps learners for discovering language errors; he also adds that children have the capacity of discovering the rules of a language system even with a limited cognitive knowledge, which is defined as language acquisition device. Usually this ability is presented during a period of time in which they achieve different levels of vocabulary, creativity, social grace, and an unconscious structure of the language

spoken. The children language is full of ungrammatical intake, but it is corrected by the environment, where he or she begins to adapt the new knowledge, and create his or her own grammatical rules.

Chomsky (as cited in Lyghtdown & Spada, 1994), calls the child's innate endowment as universal grammar, which is considerate as a group of all the common rules that exist in all languages, in this way learners acquire a second language in a similar way as they acquire their mother tongue, that means that they acquire a second language with the same variations and principles of the L1, in that process could be some variations which are proposed by the environment, but according to Chomsky (1959) children develop a language as the same way as they learn to walk", that means that the way how children learn is a natural process associated with aspects that exist in the environment, on the other hand, about the use of the principles and evaluations that we mentioned before, they can exist in a particular language that the students want to learn.

In short, we can affirm that all those theories are related with language teaching and learning processes and for this reason we consider that they can be useful for this project.

EVALUATION

Testing and assessment

Before presenting our topic of research, it is necessary to know some approaches about testing and also some important concepts like: what a test is, the characteristics and principles of a good test, and the difference among testing, assessment and evaluation.

In order to know why testing is important in a second language learning, it is necessary to study the approaches which contributed to the concept of testing that is known today.

In 1950 testing just focused on semantic, phonological, lexical, and grammatical aspects between two languages, in 1970 and the 80's the communicative part of the language was integrated to testing and for this reason, nowadays testing focuses on authentic, valid instruments and tries to submerge students into the real life. This assumptions underscores two approaches of language testing: DISCRETE POINT TESTS: This theory says that the language should be assessed into parts of the language and into some skills of the language: reading, writing, speaking and listening, phonology, graphology, morphology, lexicology, syntax, and discourse, so a test contained the four skills and the linguistic aspects which were changed due to the concepts of communicative competence, authenticity, and context, according to Collier (1979) in that time testing

took into account these aspects and in this way it achieve a big change and it was necessary to create integrative tests where the communicative part and also the linguistic features of the language were integrated, so according to (Cziko 1982, & Savignon, 1982) there were two ways of integrate tests: A cloze test: It is a reading passage where students supply words that fit into blanks, according to Oller (1979) those cloze tests results are good measures of overall proficiency because they require a number of abilities like vocabulary, grammatical structure, discourse structure, reading skills, and some other strategies. Another way for integrating tests is the dictation: In it, learners listen a passage of 100 to 150 words and they have to write what they hear using correct spelling, by this way the listening contains long pauses with the final one to give students the chance to check what they wrote. Supporters say that this is an integrative test because it requires of careful listening, writing production, memory etc. However, there were some people who did not agree with these kinds of hypotheses about testing, and for this reason, that people created a new position, it was originated by a group of students in Brazil and the Philippines, Farhady (1982) said that the assessment of the ESL depends on factors of the learners such us their native country, social status etc. By the mid-1980s, the contradictions about testing followed, and in this time, a

group of designers thought that a good test should change its view towards communicative tasks, according to Bachman and Palmer (1996, p.p 9), "the principles of language testing the need for a correspondence between language test performance and language use", so a good test should be directed towards real situations taken from real contexts. On the other hand, Canal and Swain (1980) and Bachman (1990) proposed other model where all the elements especially the pragmatic and the strategic abilities are the central points to the language testing performance. As we can see, there are some important approaches which have treated to determine the characteristics and the types of tests in order to make evaluation a successful process.

Now, it is also important to talk about the types of tests, in this way we can find many alternatives which are: Multi-skill proficiency test: it determines a general ability level; Quiz: it focus on recognizing correct use of definite articles, it measures specific knowledge; Classroom based short answer essay test: it may earn the task taker a letter grade accompanied by the instructors' marginal comments; Large scale standardized test: it provides a total numerical score, a percentile rank, and perhaps some sub scores; Performance-based test sample: it takes into account the test taker's actual use of language, but from those samples, the tests administrator infers in general competence. In short, there are several kinds of

tests which belong to constructing a good test as a complex task involving both science and art.

Proficiency test

Heaton states that the main purpose of a proficiency test is to assess someone's ability on the domain of a second language in a particular task. These kinds of tests are addressed to show how well a candidate can use the language for his or her own purpose in real life. For this reason, it is not a criterion-referenced test because it does not try to compare several candidates' results, but to show the performance itself.

Now we are going to talk about some basic concepts about testing: First of all according to Brown (2000), test is a way to know the abilities of the students in different fields of the language, it is also a set of techniques, procedures and specially a test is a method which can be developed of many ways, these ways are explicit and structured: " multiple choice questions with prescribed correct answers; a writing prompt with scoring rubric; an oral interview based on a question script and a checklist of expected responses to be filled in by the administrator, these are ways of measuring abilities, in this way we as students should ask ourselves if a test really measures individual abilities? Maybe this question contains a series of aspects, first of all a test should contain a series of principles which are: **RELIABILITY**, According to Brown

(2000), a test is reliable when it is consistent, that means that if a teacher applies a test to the same group of students different times, the results should be in a similar range, however there are some factors that affect the reliability of a test, these factors are the emotional states of teachers and students at the moment of elaborating and taking a test, another aspect is that the tests contain errors or it does not give clear directions, in this way the reliability is applied to teachers and students. Also Bachman (1990) states that it is important to examine the reliability of tests before they are applied, thus we can identify the possible errors in the tests and analyze how they can affect the test score. Another principle is **VALIDITY**, according to (Gronlund, 1998, p.226. quoted in Brown 2001) this principle is " the extent to which references made from assessment results are appropriate, meaningful, and useful in terms of the purpose of the assessment. In this way, as a general term, does the test measure what is supposed to? Having specified the aims of a test at the outset, the purpose of validation is to find out if in fact it meets that aims. Another important principle is the **AUTHENTICITY**, according to (Bachman and Palmer. 1993. p 23 quoted in Brown(2001), this principle is " the degree of correspondence of the characteristics of a given language test task to the features of a target language task. It means that a test should have some

characteristics like: the language used in it which should be as natural as possible, the items should have a relation and not be isolated parts of the test, the topics and the content in the test should help the student to learn the language, it is also important that the test contains an organization of the items studied, and finally an authentic test should present tasks in relation with real contexts. A test must also be **PRACTICAL**, Bachman (1990) claims that before applying a test we need to take into account what we need to have a successful test, it means that it has to be well organized considering the cost and the time, not only when preparing and taking the exam but also scoring it.

On the other hand, it is important to know the difference among testing, assessment and evaluation.

Testing is a process which tries to measure many aspects of the knowledge, but in the case of the language learning it is known as language testing. This type of testing tends to measure those aspects that take part in the process of learning. Tests can help students, teachers, and administrators in order to know the progress that they have obtained during the process of teaching. It is important to affirm that the language testing today shows an interest in teaching communication and learner's interaction in the society. Tests make a diagnostic of characteristics; they confirm what each person has understood according to his level or capacities to learn

something. Testing helps English teachers to know if their activities in classroom are efficient or not, and if they are doing or applying a correct syllabus. Tests also help students to know what they know about the language, and it is used as a tool which makes a kind of pressure for studying the language, and in this way learners develop their capacities and abilities.

According to Brown (2000), **ASSESSMENT** is a very wide process of measurement which is always present not only in classrooms but also in every place where a person is evaluating or correcting any error in the use of the English language. Assessment is not usually planned, for that reason we can find two types of assessment, formal and informal assessment. The latter refers, as it was mentioned before, to not planned activities or common and daily situations where the teacher can realize how learners use the language, and at the same time giving some impromptu feedback. On the other hand formal assessment is a procedure which is planned, designed and systematic and it has strict directions. This procedure is generally used in a long-term period and it is not necessarily time-constrained or arranged for a specific class.

Assessment is also divided in summative and formative assessment. The first one is probably the most common one because it is applied in most schools and universities, in this process the learner's knowledge is measured as a

percentage and tests are used as an individual and summative tool of teaching and learning process. Formative assessment provides learners an environment where their competence is being formed and not judged, this type of assessment is normally found in informal assessment than in the formal one.

According to Pilliner (as cited in Bachman 1995) there are two types of tests, one is "subjective" tests and the other is "objective" tests, these are distinguished because of the scoring procedure.

In order to create or develop a test; teachers or test developers use relevant information as curriculum content or theory of language, so that it determines subjectively the content to be covered in such test. On the other hand, tests writers make subjective decisions about how to construct test items in a successful way.

When Students take a test, they also have to make subjective decisions about how to answer the questions; such questions can be essays or Multiple-choice.

An objective test predetermined criteria helps teachers or scorers in order to give a precise score or grade on the students' correctness. On the contrary, in a subjective test the scorer has to give a judgment base on interpretations of the scoring criteria.

Some examples of objective tests are Multiple-choice questions, cloze tests and dictation can be objectively scored when test takers write a correct word or exact word.

There are two kind of formal tests: NON REFERENCE AND CRITERION TESTS, in the first one, the scores are known by the test-takers, for example TOFL, SAT etc. The second one, the form of the grades which may be of little concern as long as the instrument assesses appropriate objectives through a process which is called **TESTING**, and the broadest term which is carried out for decision - making purposes, it is **EVALUATION**, (Brown 2000).

Some authors have given concepts about what evaluation is and its importance, Weiss (cited by Bachman 1990) for example defines this term as a system of collecting information with a specific purposes. Thus, when it is necessary to make decisions in any situation, it does not only depend on the ability of the decision maker but also on the results you can get from a evaluation, for this reason the collected data is to be relevant and reliable. In this way Evaluation does not involve testing, because tests describe purposes and are used as a tool for educational and pedagogical purposes like revising a topic or material thought. Tests can only be useful for evaluation when its results are the bases of any decision.

On the other hand and talking about our project of investigation which is testing, its types and

characteristics, According to Penny (as cited in Brown, 2000) there are many important elements in the learning and the teaching process in relation to foreign languages, one of these elements is the language testing that has an important influence on its development, this influence is called "Psychometric" which is the measurement of the human mind that has a factor shows as the "factor g" or general intelligence, it is a single measurable factor. The techniques in current use in testing include correlation, regression and the analysis of variance. It is very difficult to accommodate the techniques to each learner's needs because each student has a very different level in language's knowledge, however the testing process has developed a set of different kinds of evaluations which take into account a variety of tests which include basic competences into the foreign languages especially the communicative competence proposed by Hymes (1971; 1972), in this way according to Palmer is easier to allow that students express what they know through an oral way, maybe in this sequence it is useful to develop oral abilities through communicative and integrative tests such as: interview, cloze test, dictation, free composition, etc. In this way learners will develop in a better way all their abilities in the fields of languages. It is important to know that the factor 'g' is like a base to start to measure what a learner knows but from this point to the following

progress in the knowledge only depends on many other factors like the method and the techniques that a teacher uses to improve the knowledge of his students, the materials etc. And if all those components function in a correct way the measurement of the learners according to their levels, through tests and evaluations in oral or other ways will be successful. According to Wilkins (as cited in Brown 2000) the process of evaluation in language teaching has a very important relation with the objectives that the educational system wants to achieve in order to carry students towards new perspectives to develop their abilities in order to establish and demonstrate what they know and to value how successful they have been in their learning in relation to their linguistic achievement. On the other hand, there are many ways to see the teaching process and to evaluate it, for this reason, if teachers see the language learning as the acquisition of a range of grammatical structures, the emphasis of the tests will also be on the grammar of the language, if they regard a command of the phonological contrast as important, this fact will also reflect in the form of the tests, but if the teaching is synthetic, testing will tend to be so too. That is to say, in teaching as in testing; the language is broken down into what are regarded as its components parts. The performance on such tests is not necessarily an accurate guide to what we might call the individuals "capacity to

communicate. We take this capacity to have the ability to encode a message correctly, certainly but spontaneously, fluently and appropriately. For this reason, the tests involved would inevitably be largely tests of integrated rather than isolated skills. They would also be tests of language performance. The problem of testing actual performance are well-known in the context of a ‘‘notional syllabus’’ where students can express things as concepts of time, space, movement, intentions, prohibitions, etc. The forms of testing do have a considerable influence on the manner and the content of language teaching and also allow the development of the new testing techniques that would be a contribution to teaching process.

Reasons for classroom testing

In order to know what learners have achieved after a specific course, syllabi proposes to apply a test at the end of it; it may be an achievement test or a progress one. For this purpose teachers must know what learners are supposed to be able to do.

A progress test not only should show how well learners have improved in their acquisition of a second language during the course or at the end of it, but it also tends to show how successful classes have been developed. However, if we want to measure teaching, syllabus and the progress of learners we prefer to apply an achievement test.

According to Heaton (1990 p.p 14), in certain ways an achievement test is also like a progress test but it is usually designed to cover a longer period of learning than a progress test. If we confine our test to only part of the syllabus, the contents of the test will not reflect all that has been learned.

In this chapter, we talked about the most relevant theories about second language learning, and teaching, in the same way, it was given some aspects about our project of investigation that means that one talked about testing, its types and characteristics. In the next chapter we will talk about more direct aspects of our study, more specifically about the method.

Test design

Test design is the first step a teacher faces; at this point he must think that the specifications and accuracy of them are the basis of a careful judgment and development.

When a teacher is designing a test, he must be careful when writing the *rubric*, that is to say that he must give clear instructions. Sometimes it becomes easier when those instructions are similar to some done in class or previous exercises. If this rubric is not clear enough, the results would change and in this way it would be difficult to know if the learner has achieved the purpose. This is really important because in this way students can know what they

are asked to do -which is an important aspect of validity.

(Harrison 1983).

THE RESEARCH METHOD

In this chapter we have specified some aspects of this investigation. What we are mainly concerned with here is the types of researches related with the project, the setting and the participants that worked with us, the instruments and techniques developed for the data as well as ethical issues considered in this project.

The design.

This project is related according to the qualitative research as well as the quantitative one. The fundamental reason for believing those assumptions to be true is that the first one, focuses on the educative phenomenon as well as some other aspects related with this investigation as facts and experiences of the participants involved in the educative process, this was said by Moliner (as quoted in Boyalzis, 1998); on the other hand, the quantitative research focuses on the results of the investigation, besides it tries to explain the relation of cause- effect or association- incidence which can help us to understand the causes and the incidences of our research project.

Finally, an argument for supporting the last assumptions is that one needs to focus this investigation on the evaluation processes as well on the results of it, and both researches can help us to know that.

Types of research

As we said before the research project was based on the qualitative research because it is needed to know real aspects of the project, according to Hamersley & Atkinson, 1994) through this type of research we established a permanent and closed relation with the participants that in this case are the learners and English teachers.

On the other hand, this proposal was based on the quantitative research because it was needed to describe what was happening, according to Moliner (as quoted in Bastidas, 1991) this kind of research influences on all the possible dimension of the investigation, also this research tries to collect all the possible information about setting and participants which could be useful in order to develop this project with success.

SUBJECTS OR PARTICIPANTS, SETTING AND SAMPLE

Subjects

In this proposal, we worked with 22 English learners from the ninth semester of the English and French and English and Spanish programs from the Linguistics and Languages Department which belongs to the Human Sciences School at the University of Nariño. The other participants in this paper were 6 Linguistics and Languages Department teachers who work with these programs currently. These subjects were chosen for the following reasons: First, ninth semester students have been exposed to different

assessment tasks and practices and since they are doing their practicum they have an idea of what may be appropriate in terms of assessing foreign language learners. Also, it is assumed that teachers who work with more advanced levels in the Department have a wide knowledge about L2 teaching and might apply a variety of written tasks and tests to students.

Setting

This research project took place at the Linguistics and Languages Department in the University of Nariño. The Department is in the Panamericana campus located in the north part of Pasto. This place is near a busy street where there is too much traffic and there is also a lot of pollution; these factors might have a negative influence in the learning process because the learning environment is exposed to these elements. It is also important to take into consideration that this place has excellent resources such as TV sets, DVDs, computers and a large collection of books which enhance the teaching and learning process. It was decided to work in the University setting because the available resources may facilitate the use of different assessment tasks and the application of the assessment principles.

The sample

In order to analyze how the process of written assessment is being developed, it has been taken into account that in this institution the English and French program has classes in the morning, and the English and Spanish students work at night. Both programs were included to have more input from different types of students, and also to collect more information about the assessment tasks used at the Department.

Procedure

A survey was applied to teachers and students; each one was designed with the purpose of collecting information on the nature of the writing assessment tasks and to observe if the principles of language assessment are being carried out. The results obtained from the information collected were compared to current literature on the subject of assessment in order to draw some conclusions.

Variables or factors

As we know the principal variable is the performance in the English language at the programs, this situation has related some factors, which influence and direct it towards a process in a continuous evolution. In this way, one of the factors which may influence directly on that process is testing and more specifically its form and characteristics, these two variables are influenced by other elements or intervening variables, like the emotional

state of the students when taking a test, the level of motivation to develop a test, the environment where the teaching process takes place, the preparation and training of teachers in order to design and prepare a test etc., all these aspects are involved in the educative process and for this reason they influence in a big way the development of it.

Data Collecting Techniques and Materials.

The instruments used to collect data were the surveys above described. These elements can help us have a more complete picture of the topic this research focuses on.

Once the information was collected, the data was categorized to be interpreted. For this purpose it was also taken into account the students' background in the field of language assessment.

Description of the procedure

The procedure of the research can be described as follows: First of all, it was necessary to ask for permission to the administrators in order to apply the surveys.

Two surveys were applied: One for teachers and another one for students. Thus, we could compare the information about the use of the principles of language assessment in tests which have been already applied.

The students' survey consisted of two semi-structured questions and ten close questions. Each question was

designed to find out if the principles of language assessment were evident in the tests they usually take. For some principles more than one question was formulated in order to get reliable results. The first, fifth, eighth and tenth questions provided information about practicality, the second and eleventh questions required information about validity, the third and twelfth ones asked about the principle of authenticity, the fourth seventh and twelfth questions provided information about reliability and the sixth and ninth questions provided information about washback.

On the other hand, the teachers' survey consisted of six semi-structured questions and twelve closed questions.

Each question was designed to gather information about the way or frequency teachers apply the principles of language assessment, and about the background of them.

Validation criteria

When you talk about validation criteria it is assumed that this could be affected by many external and internal factors, this is why it is worth stating at this point that the strategy for finding the validation of the instrument of collecting could be changed during the procedure of investigation.

This research is based on the principles of language assessment; the theories cited in this project are mentioned by many authors such as, Brown, Heaton, and

Lightbown and Spada. It is important to mention that learners are familiar with assessment, especially the students from the English-French program who are taking the Evaluation, Assessment and Testing course.

Pilot study

Given the fact that this study did not have as a purpose the application of any treatment, it was not necessary to apply a pilot study. Again, the focus of the study was to describe the application of assessment principles and the most objective way of doing it was observing real tests, created by real teachers.

Data analysis and interpretation

It has already been established that the analysis will be based on the information taken from surveys applied to teachers and students and from the comparison of the written tests provided by students to relevant literature on the field of assessment. From this, conclusions will be drawn on whether or not assessment principles are applied at the Linguistics and Languages Department.

To develop this research, 22 students who belong to the ninth semester from the English and French and English and Spanish programs were asked to develop the surveys. All of them had a very positive attitude towards answering our questions and cooperated with the research. In contrast, from the 6 teachers who were asked to fill in the surveys, only three of them helped the researchers with this task,

as a new total of 3 teachers. One teacher who did not demonstrate any interest in the purpose of the survey said that was never interested in investigation and would not offer any help while this investigation was being carried out. The other teachers never gave the survey back and expressed that they could not find the papers, and that it was difficult to fill it in again because of lack of time.

Ethical issues

As a first point of our investigation it is important to take into account some aspects that are very important in the development of this investigation, for instance, one has to ask for permission to all the institutions and entities that helped us in collecting information like the Language Department where it worked, to entities that lent us some services like the Internet room of the University of Nariño, the Library of this university. Also, one needed to account with the will of students, the administrators, teachers, assessors of the institution and entities that collaborated us in developing the project; it is important to say that the project did not cause any physical or psychological problem to the people or the institutions that collaborated us, to the contrary it provided a contribution to improve of any way the situation presented in the teaching and learning process in terms of evaluation specifically in the characteristics and form of tests which are the principal instrument of measuring the knowledge,

and with this project they can help them to direct their evaluative purpose towards better ways. On the other hand, as an act of confidence with the people and the institution that helped us, it referred them with other names. Finally, the last thing that needs to be said is that it is sure that our project cannot be developed only with the participation of us as investigators, to achieve the goals and the purposes that we wanted with this project, we needed the collaboration and contribution of people like assessor, students, teachers, administrators, and the institution where we worked, so in a gratitude for their services and their aids we will present the project in order to contribute to them in their normal educative development specially giving them a contribution for the improvement of its evaluation process in the kind, forms and characteristics of tests.

In conclusion, in this chapter we has talked about the kinds of investigations that took part in our project of investigation, one also specified the sample and the participants that worked during our study, on the other hand one referred to the materials and the techniques that one used to the collection of the information, and also one talked about an aspect that one considers it is very important, it is the ethical issues where we specified several points in order to give a clear vision of our purposes with our project of investigation. In short, these

were the points that one specified in this chapter about the research method.

Analysis and interpretation of data

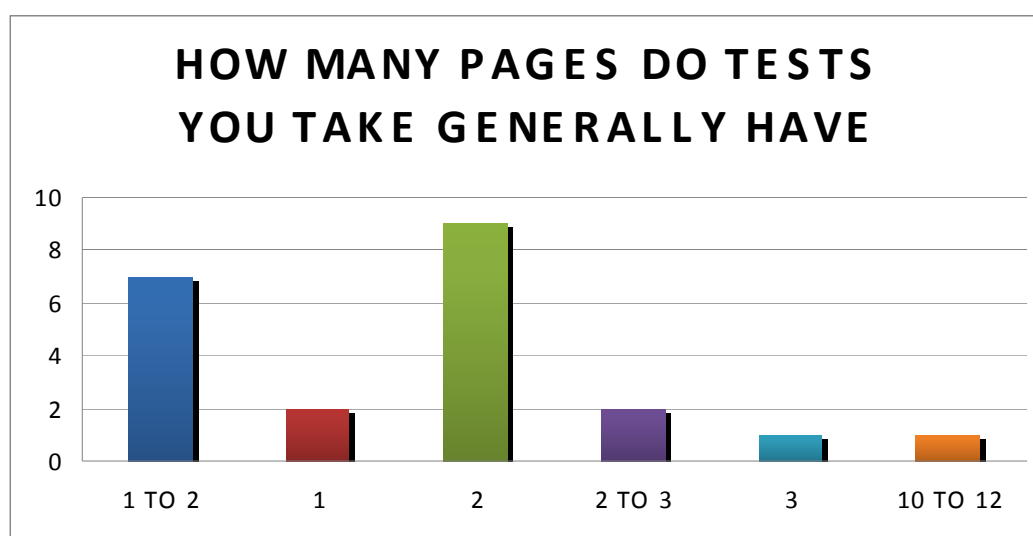
The results obtained after applying the surveys to students and teachers have been organized in the following graphs.

Students' survey

The results bellow belongs to the students' surveys.

To gather clear results about practicality students were asked some questions to know if tests are economic, easy to grade, short and clear when giving directions. The first question was, "how many pages do tests you generally take have?." The most common answers were,

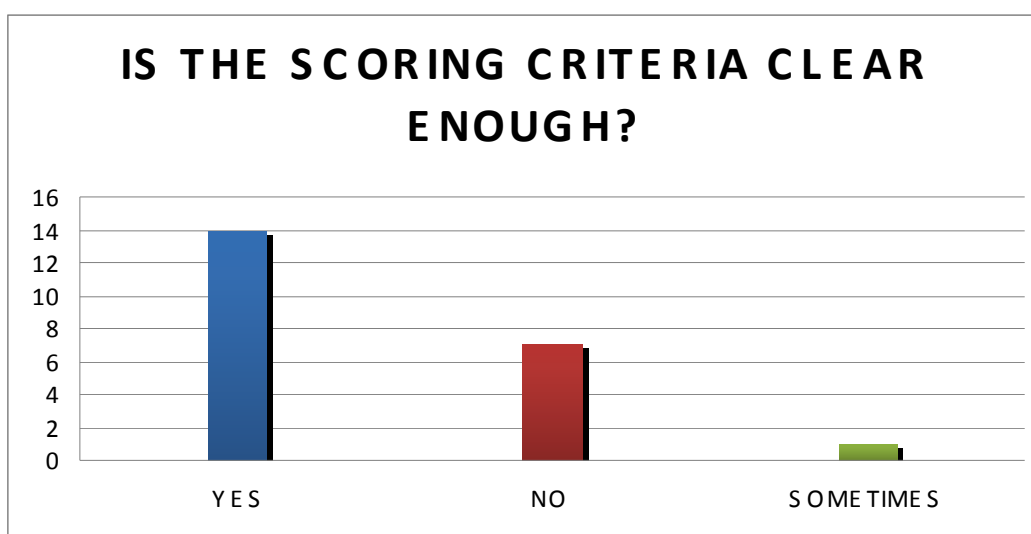
- "most exams usually have 1 or two pages"
- "sometimes 2 pages but they can be longer depending on the topic to be assessed"
- "Most of time two pages"



According to Brown, a test is effective when it is practical, which means that it does not have to be long to take and grade. Furthermore, it is practical when the examiner does not spend much money in materials.

The graph shows that 9 students considered that the tests they take are 2 pages length, and 7 of them agree the tests taken are 1 page per test. It means that most tests in the university are from 1 to 2 pages extended, which demonstrates that the principle of practicality is applied for teachers in the evaluation process, since they are as economical as possible.

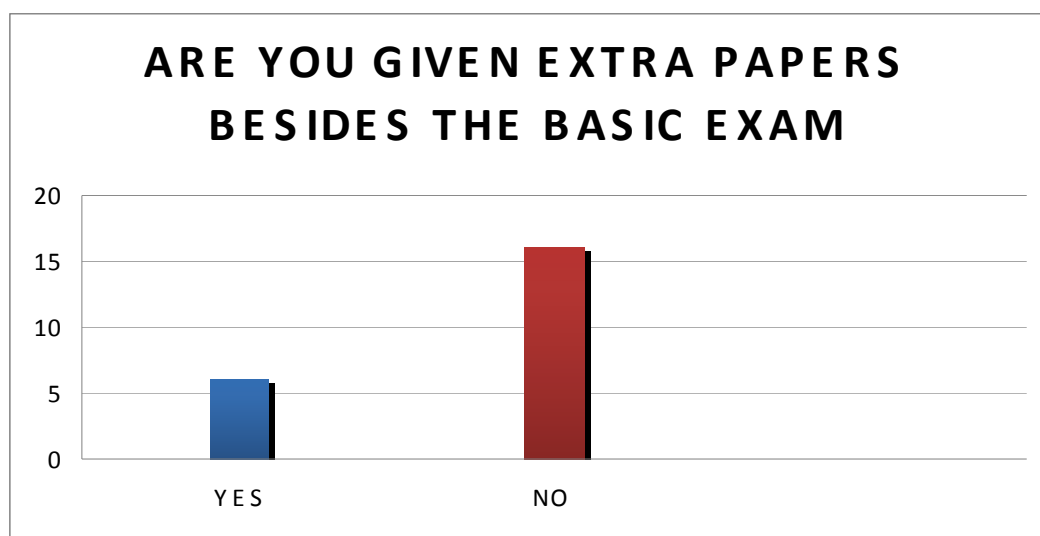
Another feature of practicality is about the score and the evaluation procedure, this questions attempted to know if they are clear and specific enough when presenting the score of each question. For this question, "Does each question of the test include the score or points?"



The graph shows that the scoring and the evaluation process is clearly carried out by the teacher. After

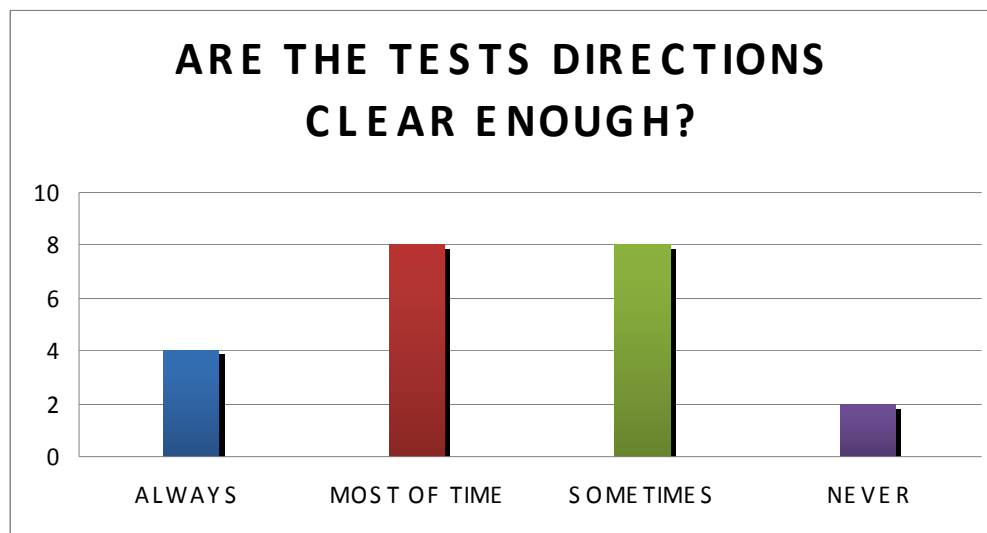
analyzing the results of this question we can say that the tests applied are practical since they give clear scoring criteria. Spite of this results some students consider that teachers do not provide such criteria beforehand.

Another aspect of practicality is if tests are economic. Students were asked, "Are you given extra papers to draft one of the points or questions of the exam?" Sixteen students answered that they use the necessary number of sheets of paper per exam, and that tests are designed to be developed without the necessity of giving extra paper.



According to this graph, 16 students stated that teachers take into account the principle of practicality in terms of length and economical issues. That is to say, tests applied in the two programs are effective and practical; they do not use more sheets of paper than the necessary ones to develop a test.

To determine if students consider the rubrics are written in a clear way, they were asked, "Are the rubrics clear enough for you when taking a test?"

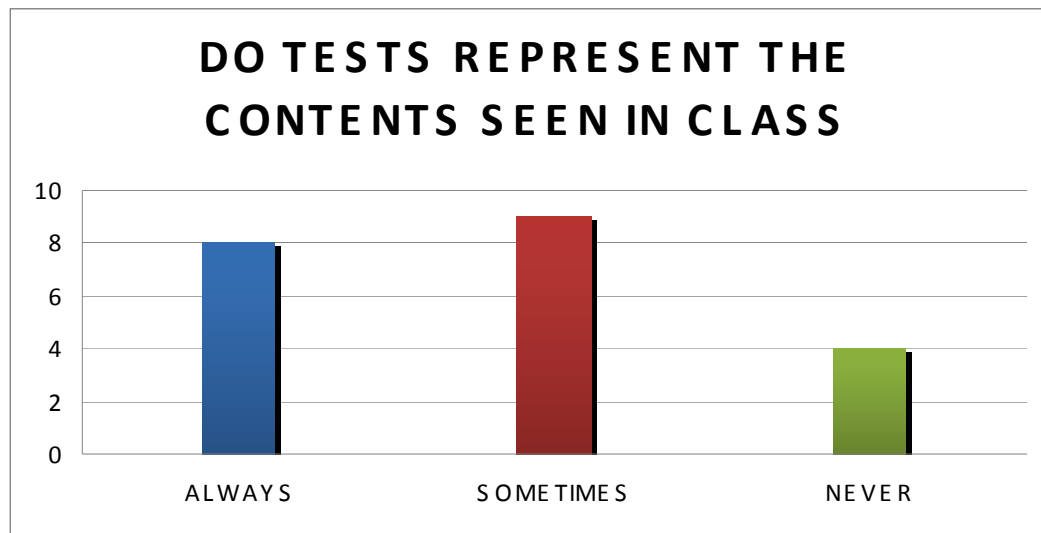


Taking into account the results the graph shows, it can be concluded that 8 students consider rubrics are sometimes clear enough and they have to ask for clarifications on those instructions. However, the same number of students said that most of time rubrics are clear to be understood. According to these results, we could conclude that exams can be confusing for some students when taking them, and that it can influence on the results exams will show.

In this respect, Harrison (1993) claims that the rubrics do not have to contain crucial explanations if the questions or directions are familiar with some exercises developed before, therefore, that kind of explanations are not important if the learners know the teacher's procedure.

In terms of validity, which is another principle of assessment, two questions were asked; the first one was formulated to know if learners consider they are assessed on what it was intended to or what the teacher said previously. The question was, "Do you think tests applied by teachers assess what they are intended to, or they agree with what the teachers is actually assessing?" For this question most students answered:

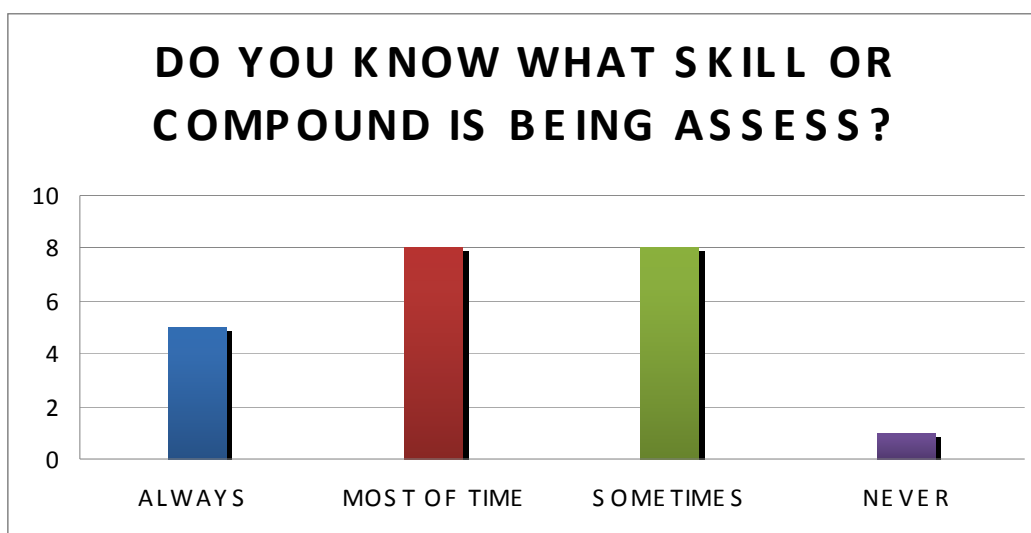
- "the exams or tests are according to the topics developed in class"
- "sometimes teachers tell students what topics will be assessed however they include in the tests different topics"
- "I think some teachers applied very good tests, but there are teachers who have no real knowledge about assessing"
- "sometimes, basically that depends on teacher; for example in assessment everything regarding test is ok but on the contrary some teachers just copy and paste some exercises from internet or books"
- "I have only one teacher who knows how to assess correctly but I think most of the exams are tricky because teachers don't know how to assess their students and they ignore how to create a good test."



As it is shown in the graph 17 students agree that the exams keep a close relation with the contents explained by the teacher and students find accurate information in the tests. On the other hand, 4 students also consider there is little knowledge on the way teachers assess them.

Brown (2004) states that a test must assess a previous knowledge in a way the evaluator can notice that the test taker has a domain in a specific skill or compound.

To know if validity is applied in the tests students take we wanted to ask them, "Do you know what skill or compound is being assessed."



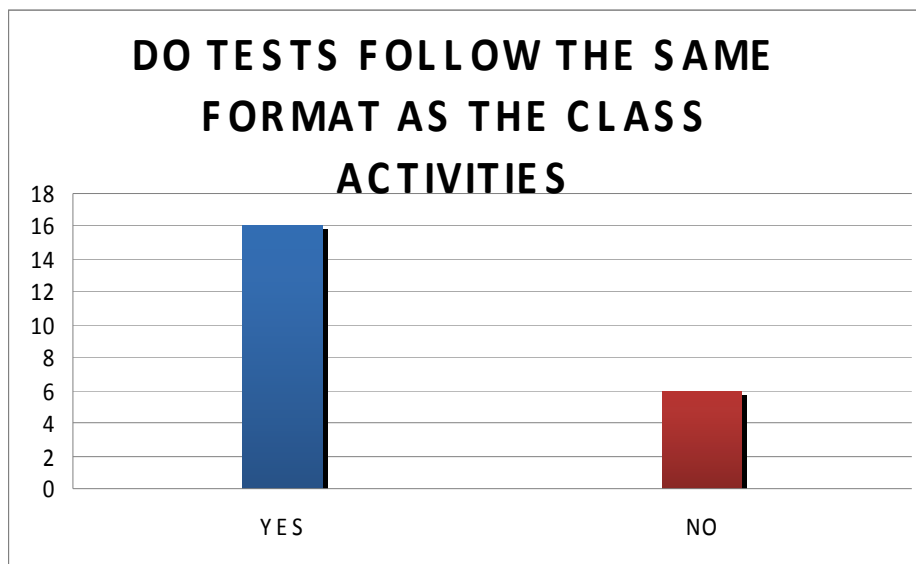
The graph shows that 8 students believe that they are assessed with a test which is not measuring their domain in the skill they were required to demonstrate. However, the same number of students said that most of time they take exams which show clearly they have to demonstrate that they can manage an acquired knowledge.

After analyzing these results it could be stated that for most students it is necessary to design test which are concise and are addressed to demonstrate what the students have achieved at the end of the course or unit of study.

Another principle of language assessment is authenticity; this principle tries to provide the student the chance to use the language in a meaningful way, with real material and as naturally as possible. To analyze if this principle is being used by teachers, two questions were formulated to gather the necessary information.

The first question asked the students, "Do tests you take follow the similar formats and contents as the tasks

done in class?." This question also has the purpose of showing if students feel familiar with the kind of questions and tasks to develop.

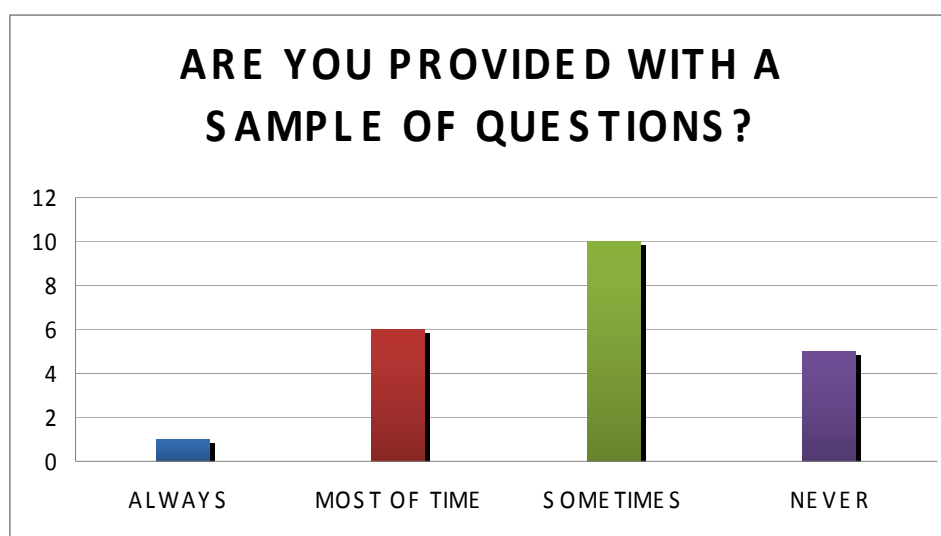


The graph shows that the target of the classes is evaluated and the objectives of the lessons are assessed through the different tests. It also shows that 16 students agree with the material teachers use to assess them, therefore they feel comfortable and they obtain good results or scores. We can realize that students not only use authentic material in class, but they also master them in any kind of test they are asked to develop.

In this respect teachers follow what Brown (2004) stated, that it to say that tests must be contextualized and as natural as possible with the use of real words and tasks.

A good test must be reliable, Therefore, the next question was asked, "Does the teacher give you a sample on

the questions or the format you may find for a future test?" The results are shown in the following graph



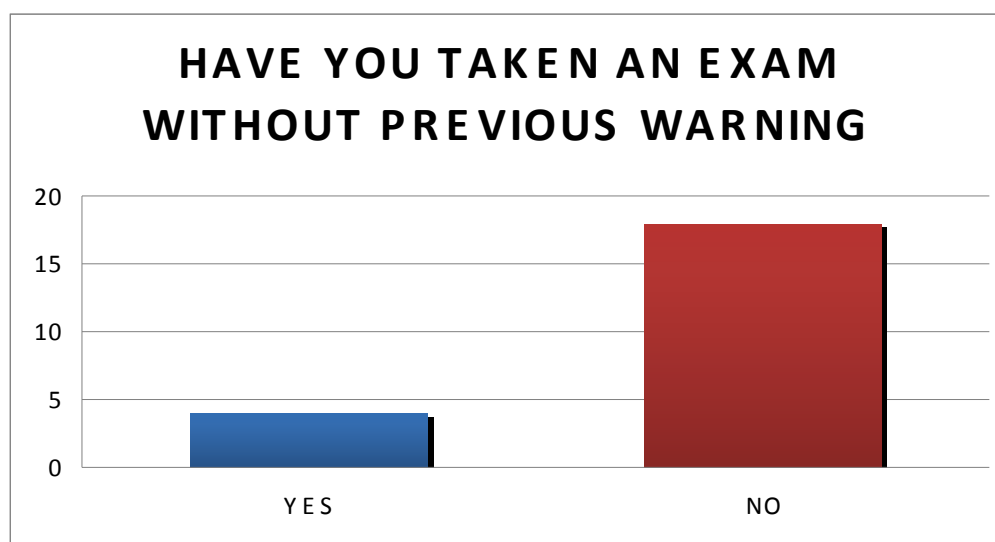
10 students stated that the teacher sometimes gives them basic questions for the future tests and that they know in advance the format and the structure of the exams. This number represents the majority of the students, for that reason we could conclude that teachers tend to apply tests which can show them trusting results. It also shows that teachers try to use alternative assessment to obtain the objectives of a specific course or unit.

According to Brown (2004), a test is reliable if it shows the same results when it is applied more than once to the same subject, in this respect if teachers give a sample of questions for a future test, learners can be prepared to answer them appropriately.

Another principle taken into account when setting a test states that the testing instrument has to measure what it is supposed to, and it has to provide trusting results.

To know if this principle, known as reliability, is being applied, three questions were formulated.

The first question intended to clarify if students have taken exams without knowing about it previously. It is important to remember that if learners do not know about an exam, it can affect on the results or scores and in this way the test cannot be reliable. Students were asked: "Have you ever taken a test without knowing the about it previously?" Students answered,

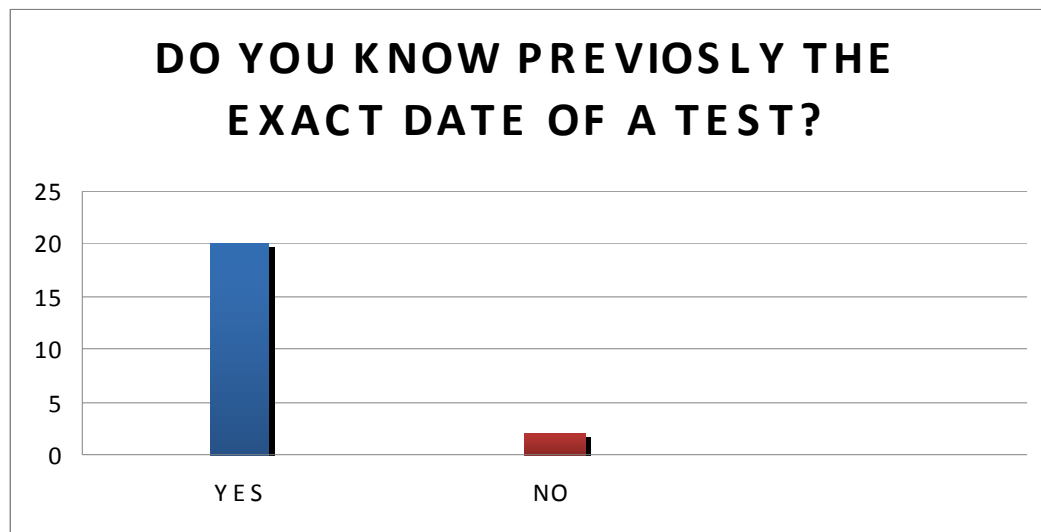


The graph shows that 18 students consider the exams are timed previously, in this way the principle of reliability is being applied and the results of those tests can be trustful, since the learner has enough time to be prepared to take a test.

In spite of a test can be timed since the very beginning of the course, there are some other factors which can influence on the reliability of the test, those factors

are illness, "a bad day", or no sleep the night before, Brown (2004).

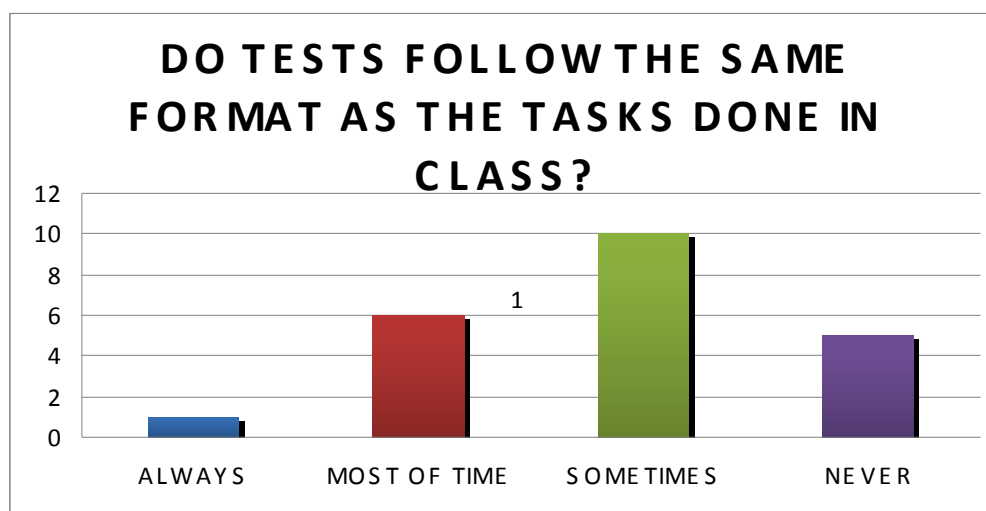
In order to confirm the principle mentioned in the previous question, students were asked: "do you know previously the exact date of a test?".



According to 21 students, it can be confirmed that the principle of reliability is being applied and their results can be trusting.

After analyzing these results, we can conclude that teachers prepare the course previously, because tests are timed and learners will be ready to take them.

A test can be more reliable if learners have a domain in the format of the test; if a test shows a format previously developed by students in class they will be comfortable. To obtain information about these students were asked: "Do test you take follow similar formats and contents as the task done in class?"

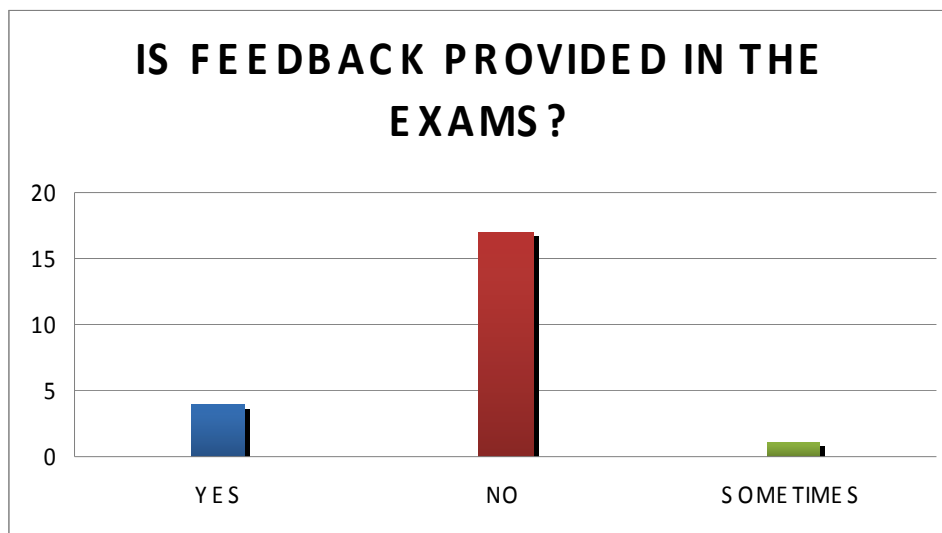


Harrison (1993, pp 20) says: "examples of tests items should usually be given as part of the rubric, but may be omitted if the tests follow a teaching procedure which the students know well. The rubric can also be understood in a wider sense to mean all the arrangements made by the tests setter enable the students to give his attention to the content rather than the form of a test."

For this reason, we considered that the students' answers show that the test they take sometimes follow this characteristic, 10 students affirm they are familiar with the format of exams. Thus, these tests can be reliable and learners will not spend time trying to understand the questions but developing the tasks.

Regarding another principle of language assessment, which is washback, students were asked about the details they find when they get a test back, those details can be suggestions, commentaries, or any other kind of additional information besides the grade or score. Students were

asked: "do the tests include commentaries, suggestions of the mistakes besides the score or grade?." Students' answers are drawn in the following graph.

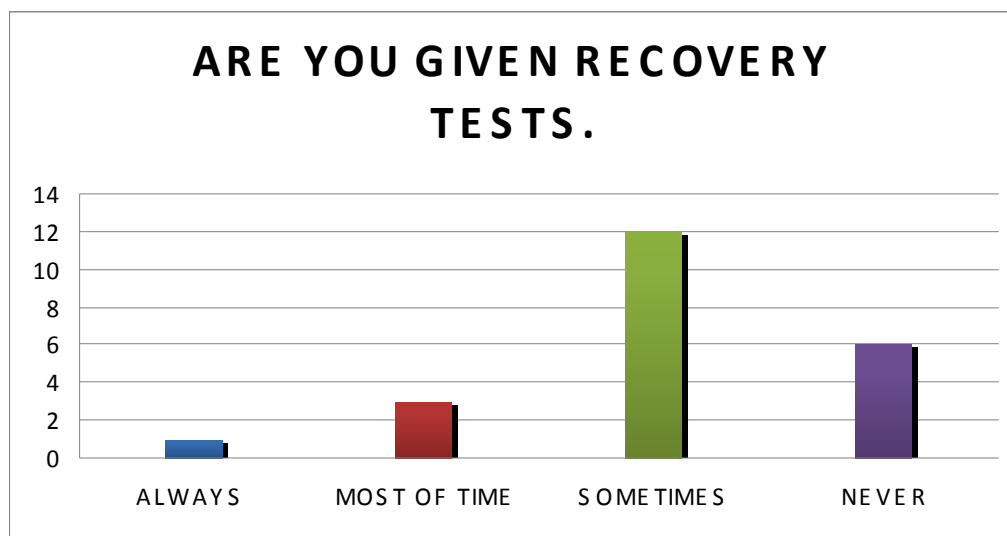


Brown (2004) claims that teachers should make suggestions or commentaries about the learners work or performance on tests, since a single grade or numerical score sometimes does not give enough information to the test takers. Students will not probably know what are their strengths and weaknesses when the tester returns the test with a representation of his or her work with a simple letter or number, as a consequence, learners will not know what the skills they have to major in are.

In this respect, Harrison (1993) says: "this will give an overall perspective and bring out any contradictions which were not noticed during the detailed work on each section."

According to the result, 17 Students affirmed that more feedback is necessary, because they state that no feedback is provided when tests are given back, and it is important to receive extra information about the tests results.

The following question asked students about the determinations of teachers when there is a low result in a test or exam in a high percentage of students. The question was: "If more than 50% of students in a class obtain low grades/results after applying a test, are you given extra activities/tests?"



According to 12 students' answers, it is believed that teachers worry and are aware of the necessity to provide extra activities to the students who have obtained not satisfactory results.

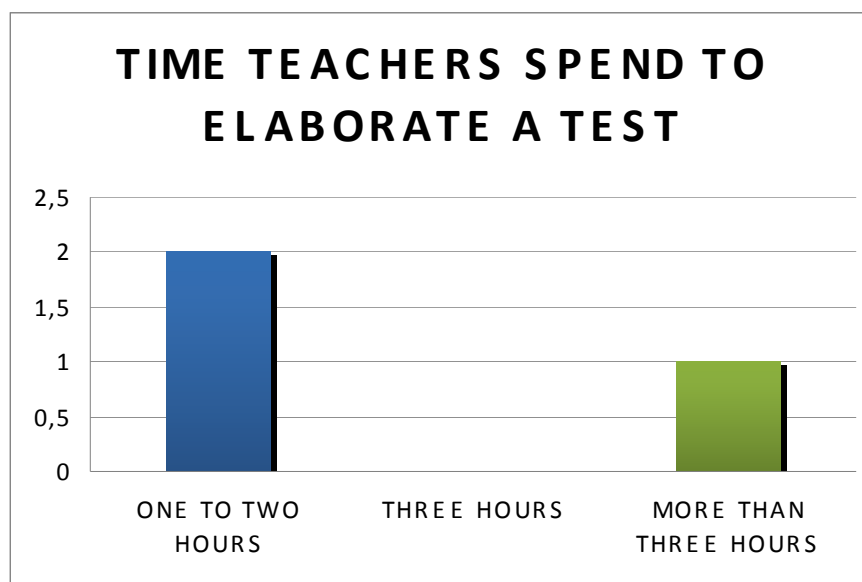
Brown (2004) states that teachers have the opportunity to provide extra sources in order to foster students' performance.

Teachers' survey

The results obtained after applying the surveys were gathered on are specified into graphs that are shown below.

Questions number one two and three asked if teachers' exams are practical in terms of economic resources and time. For question number one teachers answered the number of hours they take to elaborate an exam, for question number two teachers talked about the number of pages the exams have, and the third one shows the time teachers spend when grading a test.

The first question asked teachers: "How long does it take you to elaborate a test?" teachers answers are represented in the following graph.

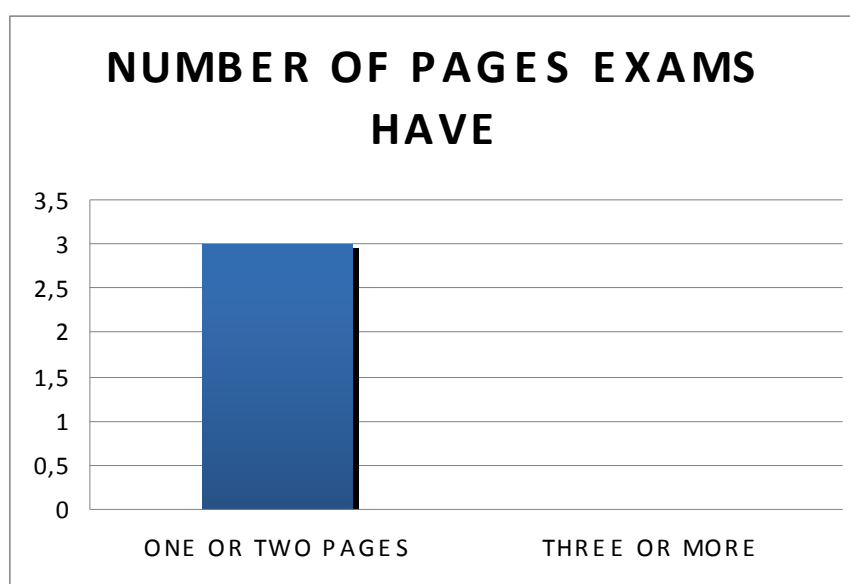


The graph demonstrates that most teachers do not spend much time elaborating the tests, and for that reason they achieve one of the characteristics of practicality.

Two teachers spend two or three hours elaborating a test, while one teacher spends a day.

Brown (2004) states that practicality also means what happens before taking the exams, which means their elaboration and design.

To find out more about practicality, we have asked: "How many pages do your tests have?."



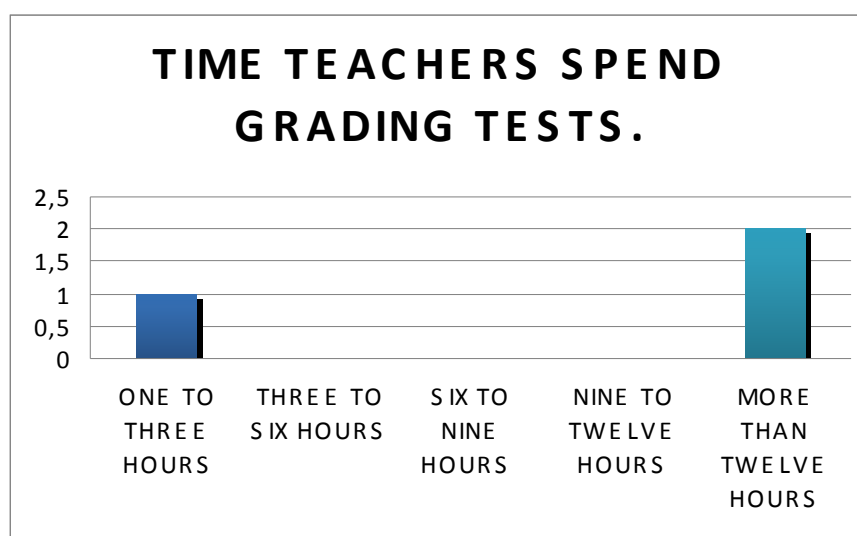
After analyzing the results, we found that 3 teachers normally use one or two pages to elaborate a test. It means that they are economic in terms of length, and that budgeted limits are considered by teachers.

The tests applied to the English and French and English and Spanish programs fulfill with practicality, because those tests are not standardized and for this reason they are supposed to be short.

Brown (2004) states that norm-referenced or criterion-referenced is something which determines the length of a

test, and as we can see in these results these are criterion-referenced tests.

Practicality does not only refer to economical issues, but it also refers to the time teachers spend creating and grading a test. For this reason we asked teachers: "How long does it take you to assess a test?."

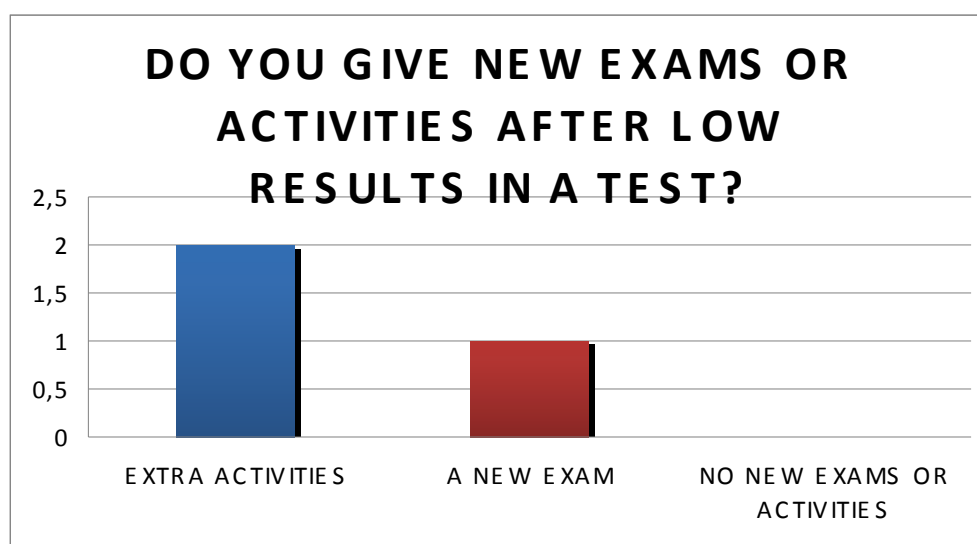


According to Brown (2004), a test is impractical if the teacher takes a lot of time to grade it and the test taker only takes a few minutes to develop it.

When analyzing the information collected, we could find that two teachers spend several hours to assess a test, while one teacher spends from one to three hours in this task. For this reason we can conclude that teachers do not take into account the time they are going to spend assessing the test applied, and it demonstrates that these tests are impractical for teachers regarding the time they spend grading.

But it is necessary to mention that factors like, the number of students in class, or the type of task learners have to develop can be a variable which can affect the practicality of a test.

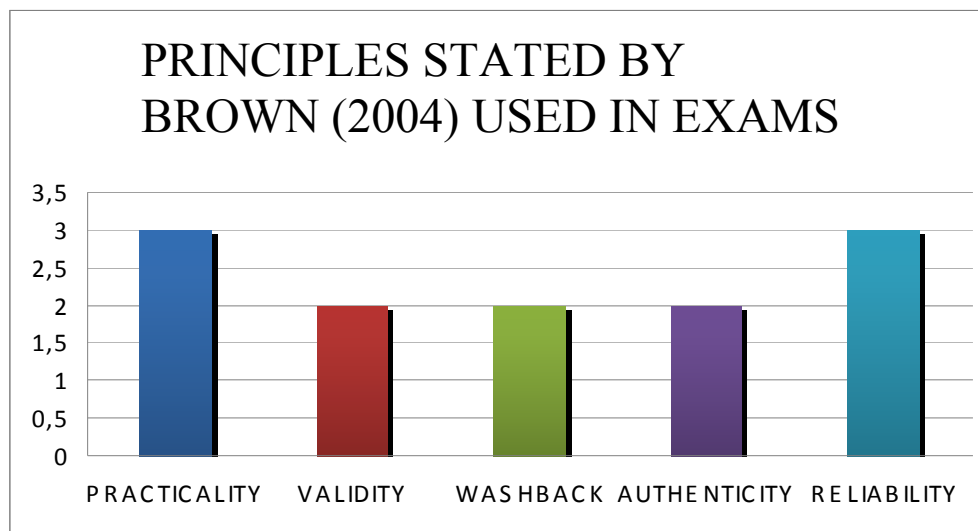
To know if validity is being applied teachers were asked about the use of extra material to reinforce the learning process of students after a high number of them have gotten low results in the test. It means that a student is given extra activities to foster the learning process so that it can be improved.



According to 2 teachers, a new test or a different activity do not replace those low grades, but they can be used to get an average. It was stated that those low - scored exams should be discussed with learners in order to provide feedback and to get information about the possible problems when taking the test.

In question number five teachers were asked about the principles they use when applying or elaborating tests.

Teachers were asked: "what principles of language assessment (Brown 2004) do you apply in your tests? The results are represented in the following graph.

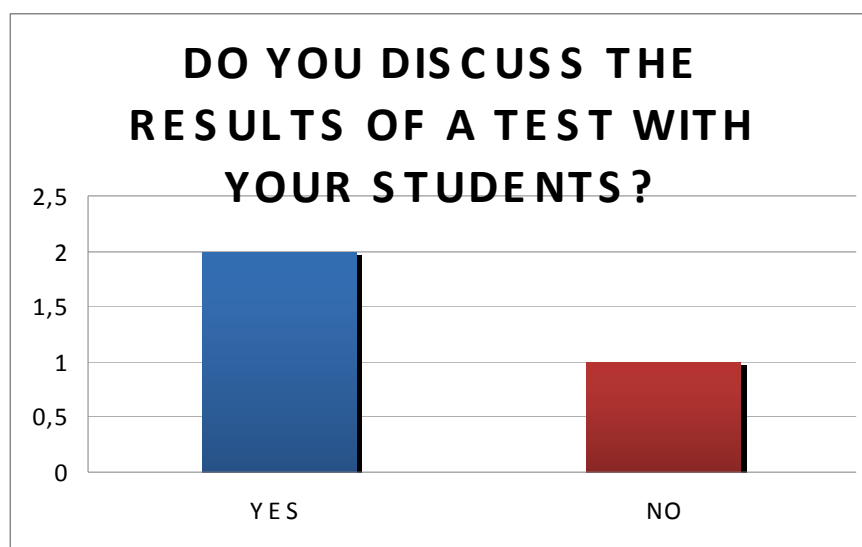


According to Brown (2004), a test will be more effective if it contains these principles.

After analyzing the data, we found that the most used principles are practicality and reliability; meanwhile the principles of validity, washback, and authenticity are not taken into account regularly.

Sometimes teachers do not realize what principle is being used, since they lack knowledge about the description of this issue.

To know if washback takes part in the process of assessment teachers were asked: "Do you discuss the results of a test with your students." Their answers are drawn in the graph below.



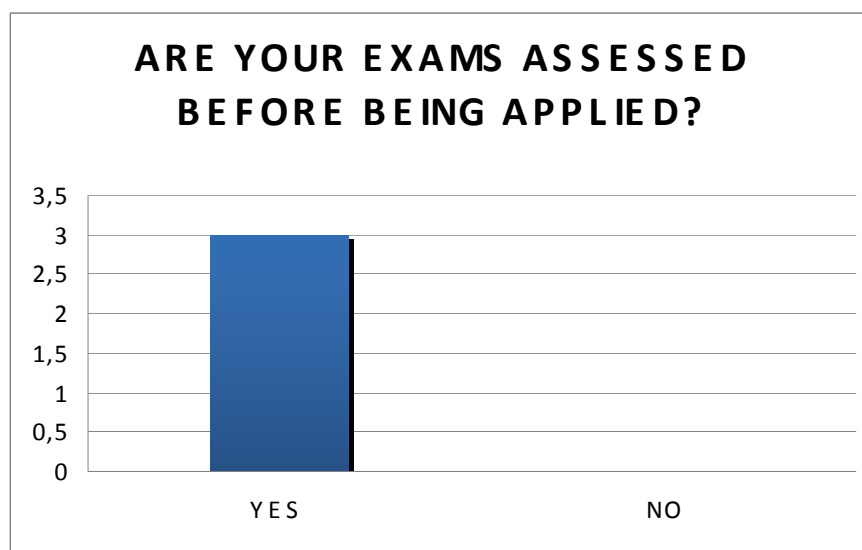
Two teachers try to solve some doubts when they give the exams back, or they try to revise the correct answers.

This type of discussion is used in order to provide a relative indicator or judgment about performance in the test itself, teachers have the possibility to foster the learning process since they stand out the strengths and provide a constructive criticism or weaknesses, besides they provide a sample of ways on how the learners can improve certain elements of performance (Brown,2004).

But it was also claimed that it is not necessary to do it unless the learners need it, this statement means that the process of assessment is not carried out in the properly, and it could be one cause of problem in the teaching process.

With the purpose of knowing if a test is valid it is important to evaluate our own tests before they are applied, after asking teachers if they follow this step.

Teachers were asked: "Do you assess your own tests before they are applied?"



Three teachers affirmed that they apply tests which have been assessed before in order to be sure they are going to evaluate what the teacher, thus, teachers elaborate reliable and valid exams.

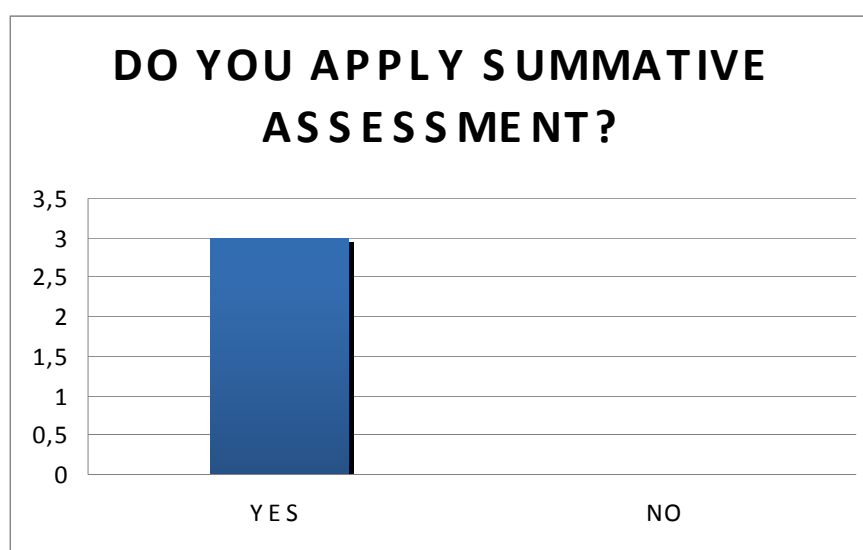
According to Brown (2004), the tests are not only reliable in their form, but they also have to at the moment of being applied, it means that the photocopied sheet is clear and that the environment and the conditions are appropriate to develop it.

These results show that tests teachers apply are an adequate to measure the learners' domain of a specific task at the end of a unit, course, or in the middle of the term.

During the teaching and learning process teachers are asked to apply summative or formative assessment.

According to Brown(2004), summative assessment normally occur at the end of a course or unit, and it demonstrates if learners have achieved a certain objective, but it does not ensure future progress in the development of the tasks.

For this type of assessment we have asked this question: "Do you apply summative assessment?"



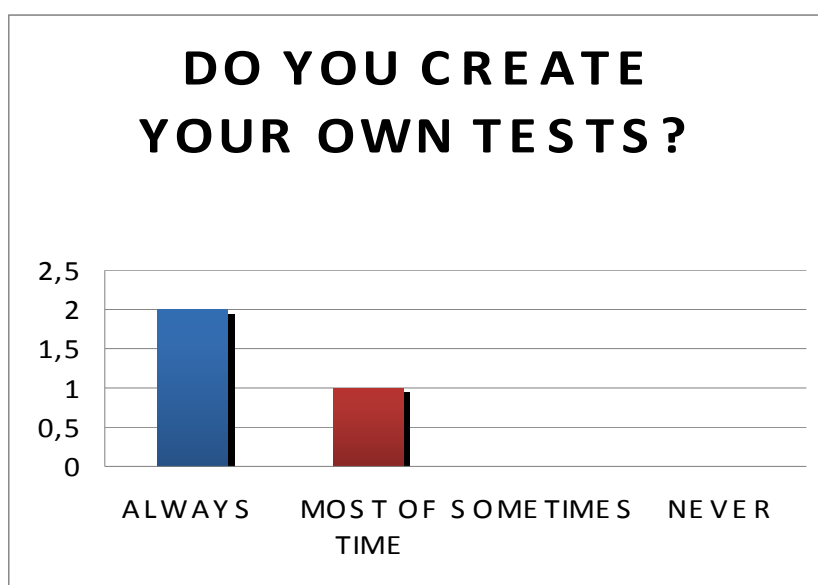
After analyzing 3 teachers' answers who claimed that they use formative assessment, it could be said that this kind of assessment is being carried out as an internal rule at work, but it does not mean teachers can also use some other kind of assessment.

Because of university policies, teachers are asked to apply summative assessment, which means that the rules of the university could influence on the objectives teachers and students want to achieve.

Brown (2004) also states that summative assessment could not contribute to the learning process, in this way learners studied just to get a good grade in a certain exam.

The following questions were set to know the frequency teachers apply some principles of language assessment.

The first question teachers were asked was. "Do you create your own tests?"

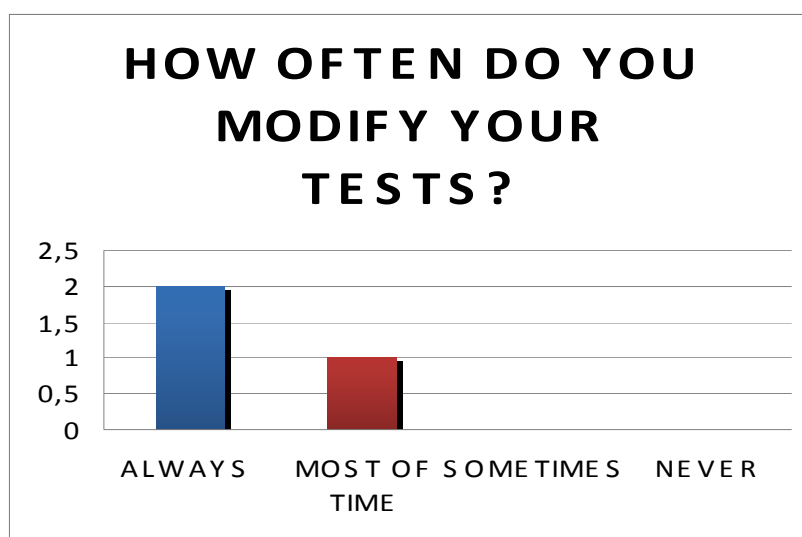


According to the teachers' surveys, two teachers prefer creating their own tests which keeps a relation with the principle of validity because the evaluators are the ones who decide which contents should be included in the exams and how these contents will be scored.

Before elaborating a test, teachers have to take into account the types of tests, achievement, diagnostic, aptitude, placement, proficiency, etc., and what they really want to measure.

Harrison (1993) says, "For diagnostic test and achievement tests the problem is less than for placement and proficiency tests because there is a body of material which is to be sampled for the test."

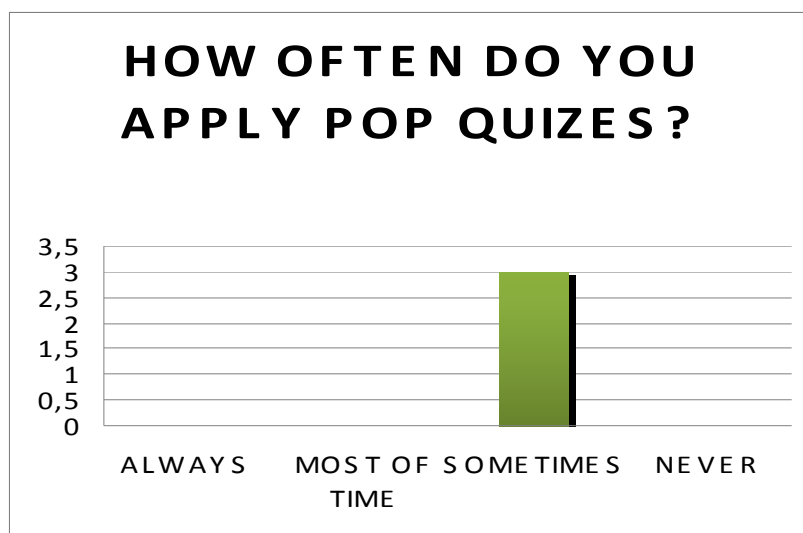
Sometimes teachers can use a test in future courses, but they should be adjusted to the new group of learners and also to the learning styles and context. For this reason teachers were asked: "How often do you modify the tests you apply to your students?"



Updating and create different questions is necessary in the learning process, it is affirmed that it is appropriate to change questions and to have a constant modification in the exams.

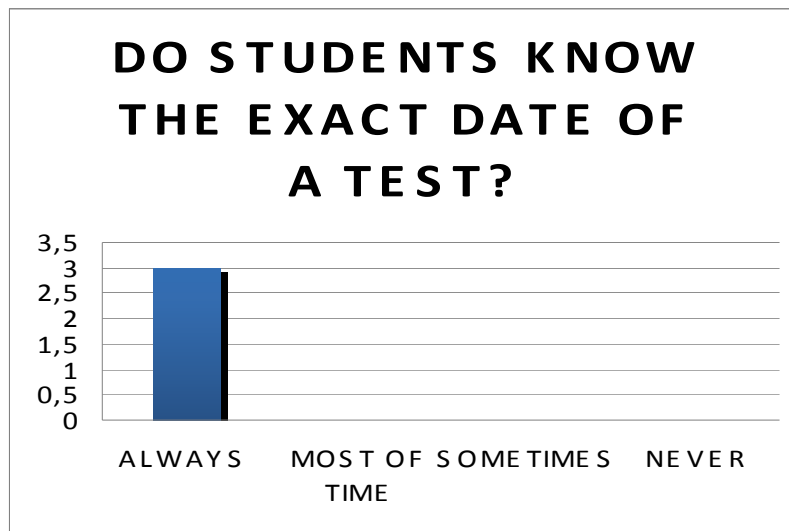
Two teachers affirmed that they always modify the exams they have applied before, while one teacher affirmed that it is done most of time. It means that most teachers worry to improve and provide better tools to learners.

A test cannot be valid or reliable if the learner does not know when it is going to be assessed because it can cause anxiety and some other reactions on him or her. For this reason we asked teachers: "How often do you apply pop quizzes?"



Three teachers answered that they sometimes apply pop quizzes. These kinds of quizzes are never planned or agreed with students, and for this reason they can provide not clear results.

Teachers were also asked: "Do students know previously the exact date of a test?"

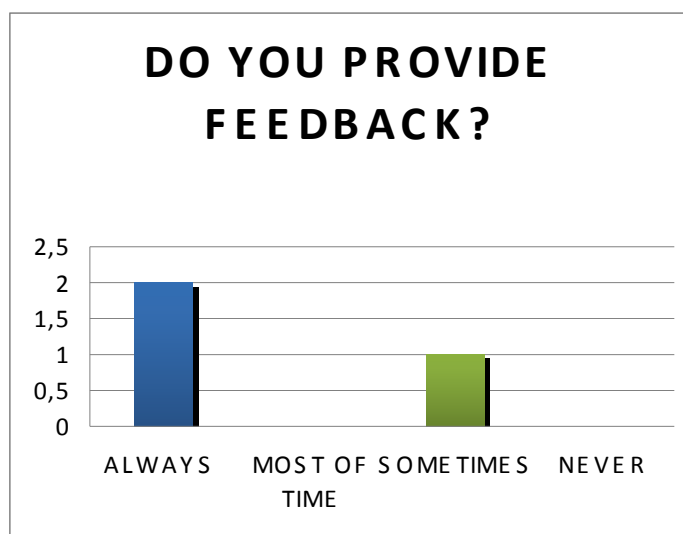


After analyzing these results, we noticed that 3 teachers said they always apply the test according to the schedule they have already planned with their students.

The previous two questions show how the principle of reliability is considered, and that teachers plan their exams and just in some cases the students are given pop quizzes which shows an appropriate schedule and a correct plan in the teaching process. However, there could be a misconception in terms of defining what a pop quiz is since there are different responses or opinions about the schedule of tests.

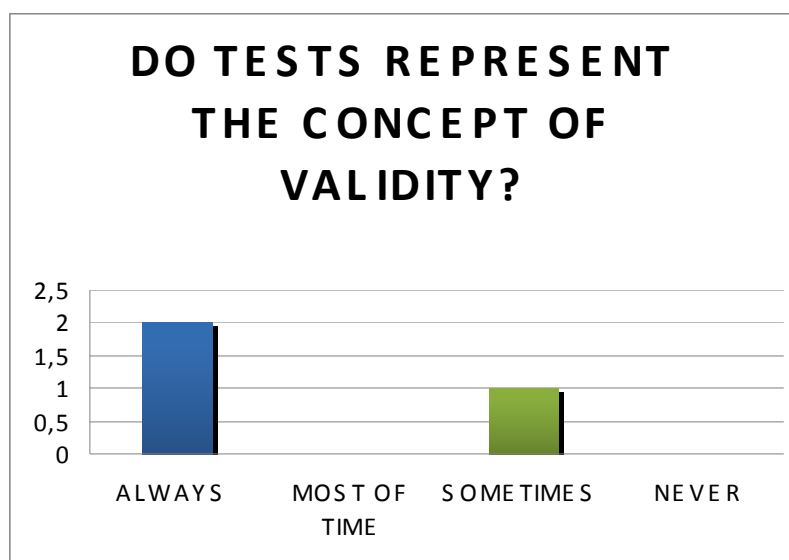
Teachers were asked if they do not only provide a number or letter which represents a grade, but if also they give extra commentaries about what has been found in the test. The gathered information was obtained through the question: "Do the tests include commentaries, suggestions of the mistakes besides the score / grade?"

For this question three teachers answered that they always provide feedback when they return the exams.



Brown affirmed that it is important to provide washback to the test takers, this way they know their mistakes and help them to improve some characteristics of their learning process, in terms of washback teachers are aware of this concept and provide feedback to their students with the purpose of getting different points of view about the evaluation process.

Teachers were asked about the application of tests on more than one group and their results to see if test are consistent and dependable. The question was: "Do tests applied in two different groups show the results same/different?"



According to 2 teachers, the tests applied to more than one group always show the same results. This means the tests are evaluating what they were designed for.

Validity, an important principle in language assessment is considered in the creation of exams and it is taken into consideration by most teachers who stated that necessity of having clear concepts in the development of tests.

Conclusions and recommendations

Conclusions

After the analysis presented previously, the following conclusions and recommendations can be drawn: First of all, there are many factors which are part of the teaching and learning process. One of those factors is assessment, which is frequently used as a tool to know if learners have achieved the goals after a course or training. However, the authors used to provide a framework for this research agree on the idea that assessment, teaching and learning should

not be disconnected elements, but assessment should be a reflection of the tasks teachers apply in the classroom if teachers want to get an accurate idea of how much students have progressed along the course. Also, having a tight connection between teaching, learning and assessment may give teachers useful information on the aspects that may need extra work, a change or what elements have been effectively learned by students and do not need to be especially reinforced.

As an important part of education, language assessment has to contain some specific characteristics in order to provide the results required by the evaluator.

Assessment has to be developed as an ongoing process. Student learning is best fostered when assessment involves a series of activities carried out along the course or semester, so that progress is controlled towards the specific course goals or objectives and the achievement of such goals.

The objectives of this study were reached successfully, and it is possible to conclude that most principles of language assessment are being applied when elaborating and applying tests.

After analyzing the results obtained, we can establish that teachers and learners agree with the features of practicality regarding the number of pages they use when setting or taking a test. That is to say, those financial

limitations are taken into account as well as the time teachers spend elaborating the tests.

Taking into account that the process of scoring is carried out only by the teacher we could noticed that they do not realize the time they are going to spend when grading tests. For this reason, this is the only characteristic which can demonstrate that practicality is not carried out appropriately.

Recommendations

Regarding that washback influence on the teaching and learning process, we tried to find out how teachers and students notice this process when evaluating.

Washback is probably one of the principles teachers need to remember to use in a wider way, since some recommendations are needed by students.

Firstly, there is a possible misconception about what washback is and the future consequences, due to students and teachers differ on the way this principle is being carried out. As we can see in the results of the study, learners say that they do not receive any kind of information about why the teacher considers his or her exam is worth a specific number or letter given as a score. In this respect learners express they need to know what the teacher is taking into account to grade the exams and they can interpretate in a better way the issues they have some problems in order to be solved or improved.

Secondly, maybe teachers use the exams results to make some decisions about the way the course is being developed, and also about the teaching process to be considered for a future change; but they do not discuss these results with the learners, they do not even write those suggestions on the sheet of the exam, in this way they are not aware of the principle of washback and it is being developed in an uncompleted way.

The concepts of the principles of validity and reliability must be clarify, due to some exams are not applied with a previous warning. It was demonstrated that students are not prepared to face an exam they do not know anything about.

One of the objectives of this study was to find out if teachers update their exams. In this respect, we can conclude that teachers demonstrate they always try to evaluate what the learners have achieved using a new set of exams or activities. However, teachers apply some tests without telling the learners previously, and in this way the exams are not going to show the results the teacher need to analyze, because the learner will not be prepared and will be influenced by anxiety, fatigue and other physical or psychological factors. That means that the test is not valid and neither reliable.

Students affirmed they sometimes feel confused with the tasks they are asked to develop since the tests do not

have the same format than the exercises or contents developed in class, or the new instructions are given in an awkward way; in this way, this problem can affect the validity and reliability of tests. As a consequence of a wrong formulation of instructions, the learner will not be able to develop the task in a good way and the test would not achieve the goals proposed or set.

It could be said that teachers feel their exams are being developed and applied in the right way, but according to learners they should pay attention to the rubrics and the time of exams.

When applying test learners noticed that the rubrics and its format are similar or follows a task developed previously, it means that teachers provide students exams which are not difficult to develop, and they can demonstrate their domain in a specific field of the use of the target language.

Finally, it was found that the English exams applied at the University of Nariño are practical and authentic, and they are providing good results.

Recommendations

After having analyzed and confronted the collected data and the concepts provided by some authors like, Brown, Heaton, or Harrison we could give these suggestions.

Firstly, it is important that teachers attend to some seminars in order to update their concepts about assessment and also to build new ways of testing students.

Regarding the principle of validity, we suggest that English teachers should agree with learners to provide adequate washback because it is the weakness shown through this study, learners need explicit recommendations and suggestions in order to improve the learning process.

Finally, as a university policy summative assessment is carried out, but it would be useful not only to apply summative assessment but also formative; in this way educational purposes could have better results, because with the current way of assessment students are facing an exam or a grade proposed, but the learner progress is not being taken into account.

BIBLIOGRAPHY

- Bastidas, J. (2000). *Teoría y práctica de la enseñanza de ingles*. Pasto: JABA Ediciones.
- Brown, H., (2004). *Language assessment: Principles and classroom practices*. New York: Pearson Education, Inc.
- Harrison, A., (1983). *A language testing handbook*. London: Macmillan Press ltd.
- Lightbown, P., & Spada, N., (1999). *How languages are learned* (Revised edition). Oxford: Oxford University Press.
- Madsen, M.S. (1983). *Techniques in testing*. Oxford: Series Editors.
- McNamara, T., (2000). *Language testing*. Oxford: Oxford University Press: Series Editor H.G.
- Penny, U. (1987) *Testing spoken language: A handbook for oral testing techniques*. Cambridge handbooks of language teachers series. Cambridge: Cambridge University press.
- Underhill, N., (1982). *Testing spoken language, a handbook for oral test techniques*. Cambridge: Cambridge University Press.
- Wilkins, D., (1982), *A taxonomy and its relevance to foreign language*. Oxford: Oxford University Press.

APPENDIX 1

UNIVERSITY OF NARIÑO
HUMAN SCIENCE SCHOOL
LINGUISTICS AND LANGUAGES DEPARTMENT

TEACHERS' SURVEY

This survey is completely anonymous and the answers obtained through it will be used only for research purposes.

Please give sincere answers according to your background as teachers of the English language focused on your experience in assessment.

❖ For the next set of questions answer them with a full answer.

1. How long does it take you to elaborate a test?

2. How many pages do your tests contain?

3. How long does it take you to assess a test?

4. If more than 50 % of students in a class obtain low grades / results after applying a test, what is the procedure to follow, do you give extra activities/ tests?

5. What principles of language assessment (Brown 2004) do you apply in your test?

6. Do you discuss the results of a test with your students?

7. Do you assess your own test before they are applied?

8. Do you apply summative assessment?

❖ Mark with an x the option you choose as your answer.

A= ALWAYS B=MOST OF TIME C=SOMETIMES D=NEVER

A B C D

1. Do you create your own tests or you use pre-designed tests which are included in the follow up text?				
2. How often do you modify the tests you apply to your students?				
3. How often do you apply pop quizzes?				
4. Do students know previously the exact date of a test?				
5. Do the tests include commentaries, suggestions of the mistakes besides the score / grade?				
6. Do tests applied in two different groups show the results same/ different?				
7. If in a “reading comprehension test” students make grammar mistakes, do you give a low score / grade?				
8. Do tests applied in a group follow the same format and content as the syllabus of a text?				
9. As a final test, do you apply pre-designed tests or you set an alternative way to assess students?				
10. If in a “reading comprehension test” students make grammar mistakes, do you give a low score / grade?				

Additional comments:

APPENDIX 2

UNIVERSITY OF NARIÑO HUMAN SCIENCE SCHOOL LINGUISTICS AND LANGUAGES DEPARTMENT

STUDENT'S SURVEY

This survey is completely anonymous and the answers obtained through it will be used only for research purposes.

Please give sincere answers according to your background as teachers of the English language focused on your experience in assessment.

1. How many pages do tests you take generally have?

2. Do you think tests applied by teachers assess what they are intended to, or they agree with what the teacher is actually assessing? Support your answers with examples.

❖ Mark with an X the option you choose as your answer.

Yes No

3. Do tests you have taken follow similar formats and contents as the tasks done in class?		
4. Have you ever taken a test without knowing about it previously?		
5. Does each question of the test include the score / points?		
6. Does the test include commentaries, suggestions of the mistakes besides the score / grade?		
7. Do you know previously the exact date of a test?		
8. Are given extra papers to draft one of the points or questions of the exam?		

❖ Mark with an X the option you choose as your answer.

A= ALWAYS B=MOST OF TIME C=SOMETIMES D=NEVER

A B C D

9. If more than 50% of students in class obtain low grades / results after applying a test, are you given extra activities / tests?				
10. Are the rubrics clear enough for you when taking a test?				
11. Do you know what skill or compound is being assessed?				
12. Does the teacher give you a sample on the questions you may find for a future test?				

Additional comments:
