

**EFL STUDENTS' PERCEPTIONS OF SCHOOL-BASED TESTING:
LESSON LEARNT FROM A LOCAL CONTEXT**

**A THESIS SUBMITTED TO THE
GRADUATE SCHOOL OF EDUCATIONAL SCIENCES
OF
BAHÇEŞEHİR UNIVERSITY**

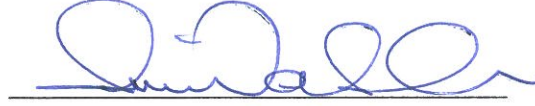
BY

Gamze TAŞLI

**IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF ARTS
IN THE DEPARTMENT OF ENGLISH LANGUAGE EDUCATION**

MAY 2016

Approval of the Graduate School of Educational Sciences



Asst. Prof. Sinem VATANARTIRAN

Director

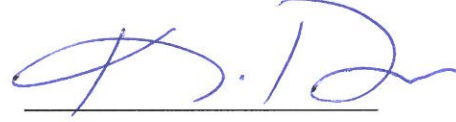
I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Arts.



Asst. Prof. Dr. Aylin Tekiner TOLU

Coordinator

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science/Arts/Doctor of Philosophy.



Asst. Prof. Kenan DİKİLİTAŞ

Supervisor

Examining Committee Members

Asst. Prof. Sinem BEZİRCİLİOĞLU

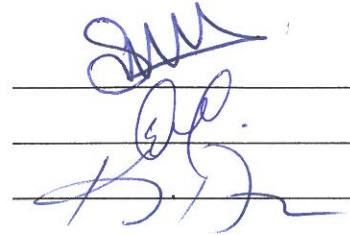
(IYTE, ELT)

Asst. Prof. Enisa MEDE

(BAU, ELT)

Asst. Prof. Kenan DİKİLİTAŞ

(BAU, ELT)



I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Gamze TAŞLI

Signature :

ABSTRACT

EFL STUDENTS' PERCEPTIONS OF SCHOOL-BASED TESTING: Lesson Learnt From a Local Context

Taşlı, Gamze

Master's Thesis, Master's Program in English Language Education

Supervisor: Asst. Prof. Kenan Dikilitaş

May 2016, 82 pages

Teaching has a potentially powerful factor in language learning and learners. Language teachers struggle to assess learner's development with formal and informal assessment. In this sense every testing situation is unique and there should be a particular role for it. In order to understand the effects of testing and assessment in language learning, it is necessary to understand the principles of testing and its relation between teaching and learning. This study aims to investigate the effects of exams on students' language learning in school of foreign language in Izmir, Turkey. The study initially investigates the learners' preferences about testing and the assessment system and analyzes the perceptions of students. Besides clarification of testing terms types of assessment and the compatibility with teaching practices were analyzed to enlighten the study. One hundred twenty-four students participated in this study. This study also aims to investigate the effects of testing from the learners' perspective. A questionnaire with the students was implemented to form the quantitative data while the qualitative data were gathered from the interview. An interview was carried out with the learners in order to identify washback effect. The findings of the study showed that learners have negative attitude towards frequency of the exams.

Keywords: Effects of Testing, Students' Perception, Washback Effects

ÖZ

YABANCI DİL OLARAK İNGİLİZCE EĞİTİMİ ALAN ÖĞRENCİLERİN OKUL ODAKLI ÖLÇMEYE DAİR ALGILARI: İçeriksel Bağlamda Edinilen Ders

Taşlı, Gamze

Yüksek Lisans Tezi, İngiliz Dili Eğitimi Yüksek Lisans Programı

Danışman: Yrd. Doç. Kenan Dikilitaş

Mayıs2016, 82

Ölçmenin dil öğrenimi ve öğrenenleri üzerinde muhtemel oranda güçlü bir etkisi vardır. Yabancı dil öğretmenleri, öğrenenlerin gelişimlerini resmi ve gayri-resmi düzeyde ölçmeye gayret ederler. Bu bağlamda, her ölçme biçimi eşsizdir ve bunun için belirli bir metot olması gerekmektedir. Dil öğreniminde ölçme ve değerlendirmenin etkilerini anlamak için önce değerlendirmenin yöntemlerini ve öğretim süreci ile ilişkisini anlamak gerekir. Bu çalışma, Türkiye'deki Yabancı Diller Yüksek okullarında, sınavların öğrencilerin dil öğrenimleri üzerine etkisini incelemeyi amaçlamıştır. Çalışma öncelikle öğrencilerin ölçme ve değerlendirme ile ilgili tercihlerini araştırmakta ve sonra da öğrencilerin ve sınav algısını analiz etmektedir. Ayrıca değerlendirme biçimleri ve öğrencinin katılımı bakımından ölçme biçiminin açıklığa kavuşturulması çalışmaya ışık tutmak için analiz edilmiştir. Çalışmaya yüz yirmi dört öğrenci katılmıştır. Çalışma, öğrencilerin bakış açısından ölçmenin etkilerini de incelemektedir. Nitel veriler görüşmeler üzerinden toplanırken nicel verileri bir araya getirmek için öğrencilere anket uygulanmıştır. Ket vurma etkisini tespit etmek için öğrenciler ile görüşmeler tatbik edilmiştir. Çalışmanın bulguları öğrencinin sınav sıklığına karşı negatif tutumunu sergilemektedir.

Anahtar Kelimeler: Ölçmenin Etkileri, Öğrencilerin Algısı, Ket Vurma Etkisi

To my parents and sweet sisters,

ACKNOWLEDGEMENTS

Primarily, I owe all my deepest gratitude to my thesis advisor, Asst. Prof. Kenan DİKİLİTAŞ for his entire support and endless guidance which also helped me in doing a lot of research as in this current study. Thank you so much for forcing me, to look at my research and my work in different ways and opening my mind. Your support was essential to my success.

Besides my advisor, I would like to thank my thesis committee members, Asst. Prof. Sinem BEZİRCİLİOĞLU and Assist. Prof. Enisa MEDE for their supports and constructive comments, but also sharing their knowledge with me.

I would also like to thank my MA lecturers Assist. Prof. Kenan DİKİLİTAŞ, Assist. Prof. Enisa MEDE, Asst. Prof. Dr. Aylin Tekiner TOLU, Assist. Prof. Yeşim Keşli DOLLAR to broaden my point of view in ELT life.

To my colleagues and close friends, special thanks to all, for their support and sharing their valuable time with me when I needed.

Finally, I would like to express my sincere love and gratitude to my beloved ones, my parents and sisters for helping me to survive all the stress from this year and not letting me give up.

TABLE OF CONTENTS

ETHICAL CONDUCT	iii
ABSTRACT.....	iv
ÖZ	v
DEDICATION	vi
ACKNOWLEDGEMENTS	vii
TABLE OF CONTENTS.....	viii
LIST OF TABLES	xi
Chapter 1:Introduction	1
1.1 Overview	1
1.2 Theoretical Framework	3
1.3 Statement of the Problem	4
1.4 Purpose of the Study	5
1.5 Research Questions	5
1.6 Significance of the Study	5
1.7 Basic Assumption	6
1.8 Definitions	6
Assessment:.....	7
Evaluation:	7
Testing:.....	7
Washback:	7
Chapter 2: Literature Review	8
2.1 Introduction	8
2.2 Definition of Terms in Testing.....	8
2.2.1 Clarification of test	8
2.2.2 Clarification of assessments	9
2.2.3 Clarification of evaluation	12
2.3 Types of Test and Testing.....	13
2.4 Types of Assessment Summative & Formative	17
2.5 The Effects of Assessment and Evaluation on Learning.....	19

2.6 Teachers' and Students' Perceptions of the Exam	20
2.7 Previous Studies on Language Assessment and Students' Perception	23
Chapter 3: Methodology	26
3.1 Overview	26
3.1.1 Research Context.....	26
3.2 Philosophical Paradigm.....	29
3.3 Research Design	29
3.4 Setting	30
3.5 Universe and Participants.....	31
3.6 Procedures	31
3.6.1 Source of Data	31
3.6.2 Data Collection Procedure.....	32
3.6.2.1 Questionnaire	32
3.6.2.2 Interview.....	33
3.6.3 Data Analysis Procedure	35
3.6.4 Trustworthiness	35
3.6.5 Limitations.....	37
Chapter 4: Results	38
4.1 Overview	38
4.2 Quantitative Section-Results	38
4.2.1 The findings of the general perceptions of test content by students.....	38
4.2.2 The findings of differences between the level of students' general perceptions of testing and assessment	41
4.3 Qualitative Section-Results	43
4.3.1 The consequence of frequent testing.....	43
4.3.2 Consistency of exams with teaching practices	51
Chapter 5: Discussion and Conclusion	54
5.1 Discussion of Findings for Research Questions.....	54
5.1.1 Discussion about the general perceptions of test content by students and different level	54

5.1.2 Discussions about the findings of differences between the levels of students' general perceptions of test	55
5.1.3 Discussion about the washback effects of classroom-based tests	57
5.2 Pedagogical Implications	58
5.3 Conclusions	56
5.4 Recommendation	60
REFERENCES	61
APPENDICES	65
A. Questionnaire Sample	66
B. Interview Sample	68
C. Curriculum Vitae	69

LIST OF TABLES

TABLES

Table 1 Module Assessment Template (2015-2016)

Table 2 Reliability Statistics

Table 3 Overview of Research Questions and Corresponding Procedures

Table 4 The General Perception of Test Content by Students

Table 5 The Differences between the Level of Students' General Perception of Tests

Table 6 Students' Perception of Testing Practices in Different Perspectives

Table 7 The Result of Students' Perception about Frequent Testing

Table 8 The Effects of Frequent Testing on Students' Progress

Table 9 Effects of Frequent Testing on Students Language Learning

Table 10 Consistency of Testing with Teaching Practices

LIST OF ABBREVIATIONS

CFC:	Cambridge First Certificate
EFL:	English as a Foreign Language
ISC:	International Statistical Congress
TOEFL:	Test of English as Foreign Language

Chapter 1:

Introduction

1.1 Overview

Educational assessment has a long and variable history. This is an ongoing process that shows the development of learners via observation and direct feedback from the learners. On the other hand assessment involves providing benchmark that helps learners see how they are doing and how they can improve themselves in language learning. A traditional examination was developed for several centuries, around 1660 when the Isaac Newton attended to the college (McArthur, 1983, p.2). There were no examinations on those days whereas lecturers decided to improve their curriculum and created their own written exams on different topics to test learners' perceptions. On the other hand, a new method of educational measurement is acquired in a review of the history of statistics. Unfortunately, the chosen work suffered a decline as the old teachers die and the task of statistic narrow down. Members of an International Statistical Congress had an attempt to find solution for the confusion. By the help of those members of ISC (International Statistical Congress) the basis for technical developments in educational statistics were formed in the 20th century. There were some efforts devoted to educational statistics such as; tabulation, averages, frequencies and variability, but at that time tabulations were mostly accepted. Many good experimental designs were developed by Galton under the leadership of Charles Darwin's studies. A number of statistical tools were improved and the first attempted for this field was measuring the characteristics of individual differences following the Pearson's chi-square test (1900), and Student's t-test (1908). There was an equitable degree of public discouragement about educational testing. Approaches to language testing began with the thought of Behaviorism (1950). According to the researchers, testing concerned with features of language like the phonological, grammatical and lexical contrasts between two languages. In this way, the first contribution to what is now recognized as classical tests theory occurred. A large-scale investigation was conducted on the use of test to decide overall class and school performance, recognizing individual skill levels and

individual differences, attitudes and personality traits. During the first decade of the 20th century there were several measurement successes in different fields. Bachman (2000) claimed that ‘current thinking in applied linguistics about the nature of language ability and language use’ has led to the development of new tests. One of the remarkable aspects of studies of assessment in recent years has been the interactions between assessment and classroom learning. According to Alderson and Banarjee (2002) ‘an understanding of what language is and what it takes to learn and use language’ is central to language testing. Bachman and Palmer expressed the importance of assessment in their studies that language assessments are primarily used to promote beneficial consequences for the stakeholders, or the individuals, programs or societies that will be affected by the assessments.

Taking all these into consideration, there is a noticeable need to conduct testing and assessment studies concerning the effects of language learning since it plays a crucial role in language learners and educators’ performance. The present study also tries to explore the washback effects on English test on the teaching of English. As Spolsky stated, “given the crucial role of the examination, the manipulation and reform of the English texts is believed to achieve beneficial washback and improve education.” (Weir, 1990; Spolsky,1996). Therefore the present study aims to evaluate how effective testing and assessment in language learning and to investigate various types of test formats and ways of testing of an English Preparatory Program at Preparatory School in Izmir.

The results of this study are expected to provide comprehensive information regarding the effectiveness of the testing system for further improvements. Finally, another crucial phase of this study is to enable for language learners to have a better understanding of the role that assessment in Preparatory School in teaching and language learning. In this sense, the result of this current study may be intriguing and suggestive for the other universities in understanding the efficiency of their testing and assessment units.

1.2 Theoretical Framework

As it is known, testing has a great influence on teaching and learning. As Alderson and Wall (1993) emphasizes that tests are believed to be effective indicators in classroom atmosphere. The idea in which assessment and evaluation effects teaching and learning is called “washback”. Hughes (2003) also mentions this term in his research as “backwash”. Spolsky (1994) on the other hand stated that ‘if the primary goal of the assessment is the control of curricula, the theory of backwash concerns with the surprising reactions of testing rather than expected effects’. There were numerous definitions of washback that the researchers share. Some of the researchers took a narrow focus on teachers and learners in classroom settings, while others included reference to tests' influences on educational systems. In this sense, Shohamy (1992, p. 513) also focuses on washback in terms of language learners as examinee when she describes "the utilization of external language tests to affect and drive foreign language learning in the school context". In order to understand the definition of washback Shohamy (1993,p.4) expressed four key definitions ;

- *Washback* effect refers to the impact that tests have on teaching and learning.
- Measurement driven instruction refers to the notion that tests should drive learning.
- Curriculum alignment focuses on the connection between testing and the teaching syllabus.
- Systemic validity implies the integration of tests into the educational system and the need to demonstrate that the introduction of a new test can improve learning.

These ideas are reconsidered in Shohamy, Donitsa-Schmidt, and Ferman (1996, p.298). In most cases washback can be analyzed in two significant types: positive and negative whether it has beneficial and harmful impact on educational practice. Anderson and Wall (1993) defined positive washback like “Teachers and learners will be motivated to accomplish their teaching and learning goals”. According to his experience as both a test developer and teacher educator Hughes (2003, p.53-56) summarized the beneficial backwash in educational system in the following list;

1. You should be clear of the abilities you want to encourage
2. Give more examples.
3. Adopt direct testing.
4. Objectives of the exams are important in testing.
5. Make sure whether the learners comprehend the test or not
6. Teachers' need to be guided when necessary.

In the classroom setting, Cheng (2005) declared negative washback as, 'The tests may well fail to generate a comparison between the learning methods and/or the course objectives to which they should be relevant. In educational settings, Hughes (2003) asserted harmful backwash 'if the content of the test and test methods demonstrate differences it refers to harmful backwash. To show the effect of testing on teaching and learning Washback is the adoptable method for the present study.

1.3 Statement of the Problem

Language acquisition or language learning is not an easy process. It requires a lot of devoted struggle to learn a language. In this sense, testing can be defined as a method of assessment and improvement of the learners. At Preparatory School, there are four terms in an academic year and two-month long in each term. In each term, there are midterm examinations and final examination, online speaking exam, writing and listening assessments are applied besides listening and on-line speaking assessments to assess and measure learners' language learning. The program in language learning is skill-based that all four skills are taught separately which are listening and speaking skills, reading and vocabulary skills, grammar, and writing skills. All of these skills are tested and evaluated via previously mentioned midterm and final examinations, listening and writing assessments, and online speaking exams. Although there are a lot of assessment types in the Preparatory School, both learners and language teachers have some concerns about the frequency of the assessment and evaluation system. Based on class observations and discussions with colleagues, the researcher herself found out some negative perceptions towards the frequency of testing.

Considering this problem, the researcher has decided to carry out a study through a comprehensive study about the effects of testing and assessment in language learning regarding the EFL learners.

1.4 Purpose of the Study

In respect to the above stated issues, the purpose of the current study is to investigate the learners' preferences about testing and assessment. The perceptions of the learners towards compatibility of the assessment system and class activities will be analyzed to identify and determine the matches and mismatches between testing and teaching in the School of Foreign Languages at a foundation university in Izmir, Turkey.

1.5 Research Questions

This study is conducted to find out how to improve the existing testing program, to identify the perceptions of the learners, instructors and testing unit members towards the system focusing on definitions of testing and assessment elements, types of tests and testing, techniques of testing, types of assessment, objectives, teachers involvement in testing, and finally strengths and weaknesses of the program. With this aim, the following questions were addressed:

1. What are the general perceptions of test content by students?
2. What are the general perceptions of tests by different levels of students?
3. How do students perceive the testing practices in terms of frequency, compatibility, clarity, length, complexity, and variety?
4. What are the students' perceptions of frequent testing?
5. What are the students' perceptions of consistency of testing with teaching practices?

1.6 Significance of the Study

This study has some assumptions to clarify the definition of testing; distinguish the testing types related to summative and formative assessment and to see the effects of testing and assessment in language learning. By using suitable classroom assessment strategies and techniques, teachers can increase their students'

cognition and also learners have chance to see how well they have learnt the language. Evaluation over reaches students' achievements and language assessments to consider all viewpoints of teaching and learning, and to look at how educational decisions can be instructed by the results of alternative types of assessment and evaluation. On the other hand, there is not sufficient research on formative and summative assessment in Turkey. Testing and evaluation studies are all related to measurement of validity and reliability of test and testing learners' skills. The main purpose of this is to provide testing practitioners, as well as other educators, with a core study to assist in making appropriate determinations regarding the assessment of ELLs in content areas. There is a limited number of studies that assess testing and assessment both from language teachers and learner's perspective. This article will be beneficial for the instructors and examiners to develop their testing system and make their teaching more effective for their learners and examinees.

With the purpose of extensive investigation and promote the field of the Formative and Summative testing and how these types of testing affect the learners' language learning ability, this paper intends to look at testing and assessment from a different point of view, as well as offering suggestions for the future improvements of Preparatory Schools which will be beneficial for their existing programs.

1.7 Basic Assumptions

It is presumed by the researcher that the participants gave impartial and truthful responses to the questionnaire and interview. It is also assumed that the learners stand for the general characteristics of target community. Finally, in the current study, the data collection instruments are thought to be agreeable, reliable and consistent.

1.8 Definitions

In this part, terms used frequently throughout the current study is defined briefly to make sure a consistency and clarity.

Assessment: The means of collecting information to observe progress and make educational determinations if necessary (Overton, 2008).

Evaluation: Systematic assessment of the operation and/or the outcomes of a program or policy, compared to a set of explicit or implicit standards, as a means of contributing to the improvement of the program or policy (Weiss, 1998).

Testing: Refers to the process of administering a test to measure one or more concepts, usually under standardized conditions (Braun, Henry, Anil Kanjee, Eric Bettinger, & Micheal Kremer, 2006).

Washback: The effect of testing on teaching and learning is known as washback (Hughes, 1989).

Chapter 2: Literature Review

2.1 Introduction

The literature of the study gives information on testing and assessment in language learning. Definition of testing elements, types of tests and testing, summative and formative assessment types, the contribution of assessment and evaluation to the learning, teachers' involvement in testing and learners' and teachers' perception of testing are mentioned in this current study.

2.2 Definitions of Terms in Testing

The field of language testing has been rapidly growing in the last few decades, in terms of both its theoretical concerns and its direct and concrete applications. It is important to define some frequently used and/or sometimes misunderstood terms.

2.2.1 Clarification of tests. Tests are procedures for measuring ability, knowledge or performance. Testing constitute one of the fundamental bases of language teaching. As testing is in every person's life, almost every step which we take in our educational life is assessed. In this way testing can be seen as a need because every continuing process needs to be seen whether it is going well or not. As McNamara (2000) states, "language tests play a crucial role in people's lives, acting as a gateways at important transitional moments in education, in employment, and in moving from one country to another". It is important to say that tests play crucial role in deciding what to test and evaluate according to the needs of particular educational program. Testing is the system of assessing learners' ability or knowledge in a specific context. Tests are the instruments that are often cautiously designed and organized, and that have clear rubric scoring. It is necessary to understand the aim of the test and what criteria we use as they have an effect on both the test takers and the curriculum issues. Before the test preparer can even begin to plan a language test, we must establish its purpose and purposes. The initial objective of testing is to determine readiness for instructional program. Some tests are used to separate those who are prepared for an academic or training program from those who are not. Another important aim is to classify or place individuals in appropriate

language classes and should be assigned to specific sections or activities based on their current level of competence. In order to test the individual's specific strengths and weaknesses some of the tests generally consist of several short but reliable subtests measuring different language skills or components of a single broad skill. On the basis of the individual's performance, profile that will show his relative strength in the various areas tested. Another objective of testing is to measure tendency for learning in order to test future performance. At the time of testing, the examinees may have little or no knowledge of the language to be studied, and by the help of the test exam preparers can evaluate or determine learners' potential. The other essential target of testing is to measure the extent of student achievement of the instructional goal. Achievement tests are implemented to show group or individual progress towards the targeted study. In order to evaluate the effectiveness of instruction tests have a great impact on achieving the goals. Other achievement tests are used exclusively to assess the degree of success not of individuals but of the instructional program itself. Such tests are often used in research, when experimental and 'control' classes are given the same educational goals but use different materials and techniques to achieve them. Moreover, there are different types of language tests, which are carried out with specific purposes. Test preparers use those tests to obtain information about the information and they categorized test according to the kinds of information being sought. In the further study, the types of tests will be described. As it is clearly seen, a well-prepared test can enrich the several aspects in language program by making rearrangements and implementing.

2.2.2 Clarification of assessment. Learners are often assessed and tested to see how well they are and how is their improvement. Another alternative of testing student is assessment which is an ongoing process that surrounded a greater domain. It is a term, which refers to all activities undertaken by teachers and by their students in assessing themselves. By doing this, teachers provide information to be used as feedback to modify teaching and learning activities. Assessments are designed to stimulate growth, change and improvement in teaching through reflective practice. Even it was a response to a question, a comment, or a new word that the students tried out, the teacher subconsciously assesses the students' performance. In order to fulfill the needs of language learners and teachers who instruct them, assessments

can be applied to inform program evaluation and instructional design. Any evaluation of learner can be assessed and all tests are assessments. As Black and William (1998) stated, 'the choice of tasks for classroom practices and assignment is crucial. Tasks have to be confirmed through learning aims that the language teachers serve, and students can work well only if opportunities for them to communicate their evolving understanding are built into planning. Discussion, observation of activities, and marking of written work can all be used to provide those favorable circumstances, but it is then important to observe or listen carefully through which learners develop and display the state of their understanding. Thus, it is asserted that opportunities for pupils to express their understanding should be designed into the teaching, for this will demonstrate the interaction through which formative assessment assist learning'. Assessment is a process of knowledge production that is the generation of inferences concerning developed competencies. In this way, such competencies are developed and these are the potential for learner's development. As a best structured coordinated system, assessment focused on the collection of relevant evidence that can be supported through various inferences about human competencies. This is based on human judgment and perception, the evidence and inferences can be used to appraise and improve the process and effects of teaching and learning. Educational assessment will certainly need to be responsive to conceptions of human mental competence and the implicit conditions by which mental competence is acquired. In the process assessment education, become an integral part of assessment, teaching and learning which is called pedagogy. The product of assessment, teaching and learning aims are reflected in the achievement of mental competence, which references the developed abilities to perceive critically, to explore widely, to bring knowledge and technique to bear on the solution of problems to test ideas. On the other hand, to recognize and create material and abstract relationships between real and imaginary phenomena can be referred to the product of assessment.

The purpose of assessment can be categorized as an assessment of educational outcomes, which is reflected the use of evaluation and assessment for teaching and learning that reflects its use for the identification of the problem and intervention. Assessment can serve multiple purposes for education. On the one hand, some purposes require precise measurement of the status of specific

characteristics; on the other hand, other purposes need the analysis and documentation of teaching, learning, and developmental processes. In all cases, assessment instruments and procedures should not be used for other purposes other than those for which they have been designed and for which appropriate evidence has been obtained. Even though assessment is a main element of education and should be associated with both teaching and learning goals, it is not the initial or dominant means for learners' development. To be productive and in education and teaching, institutions must be laid out with clear and precise teaching and learning goals in mind and promoted in ways that let them to reach goals. Teachers' abilities must be promoted with the appropriate professional development and other resources like materials, time and technology. The best assessment can advance the acquisition of competencies if the teachers guide and authorize learners to measure their progress.

Testing and assessment environment have also some differences in themselves by different purposes and for several measurements. While tests are prepared administrative procedures, assessments are an ongoing process that encompasses a much wider domain. Tests occur at identifiable times in a curriculum and learners gather all their abilities. The teacher makes assessment on students' performance on each aspect about students. By doing tests, learners' responses are being measured and evaluated but in assessment, there are many procedures and tasks different from test to evaluate learners. Generally, tests are subsets of assessments. The concepts of testing, assessment, and diagnosis continue to be considered interchangeable by many, although they have incredibly different definitions and educational values attached to each of them (Mitchell, 1993). According to The Standards for Educational and Psychological Testing (1993), tests give only the scores and offers information related to the examinee's strengths and weaknesses in oral language and literacy. On the other hand, an assessment is a more general process of gathering data to examine. An assessment process tries to find the reason why the learners perform this way. The Standards (1999) also explain assessment 'any systematic method of obtaining information from tests or other sources used to draw inferences about characteristic of people, learners, or programs. ' Assessment has some steps to control the development of learners like interviewing and observing and it is beneficial in learning process.

2.2.3 Clarification of evaluation. As it was indicated from the very beginning of this study, there were some misunderstandings between the terms of definition in testing, assessments, and evaluation. It is a common belief that evaluation has the same meaning as testing and assessment. However testing is only one component in the evaluation process. As Germaine and Rea-Dickins (1993) declared, “Evaluation is not restricted to the context of education; it is a part of our everyday lives”. Being provided a wealth of information to use for the future direction of classroom practice, for the planning of courses, and management of learning tasks and students, evaluation has also great impact on testing process. Despite this progress, there are still many crucial problems faced by the definition of evaluation. According to House (1980), none of these problems is due to the lack of definition of evaluation. The main reason for the inadequate usage of evaluations is presumably due to the lack of tailoring of evaluations to adjust the needs of the client. It is clearly stated that evaluation looks at the core objectives, at what was accomplished, and how it was accomplished. In the evaluation process, not all evaluations serve the same purposes. Some of them serve a monitoring function and some are focusing on measurable program outcomes or evaluation findings. Stake and Schwandt (2006) adapted the essential purpose of a program evaluation can be to determine the quality of a program by planning a judgment. In educational evaluation schools needs evaluation to show effectiveness to funders and other stakeholders, and to provide a measure of performance. Furthermore, individual educators need also professional activity to attempt. They intend to review and enhance the learning they are endeavoring to facilitate. It is undeniable that, the accuracy in student evaluation will provide sound, accurate, and credible information about student and their performance.

In the process of testing the definition of assessment and evaluation also dissociate between each other. While assessment is a formative and continuing process, which collects information on the extent of learning, evaluation is a summative time process and sums up achievement at the end of the course with a grade. The focus of the measurement in assessment is process-oriented and pertains to “ How learning is going? ”. On the other hand, evaluation is product-oriented and concerns about “What has been learned?” Assessments are reflective that initially defines the goals but evaluations are prescriptive, they are externally imposed

standards. Findings and the uses of those findings in assessment are diagnostic which identifies areas of improvement, but findings in evaluation are judgmental and arrive at an overall grade. Assessments have flexible ongoing criteria, however evaluation has fixed criteria that reward success and punish the failure. Standards of measurement in assessment are absolute that tries to reach the ideal outcomes. In this sense, the measurement of evaluation is comparative. Therefore, assessment for formative purposes is formed to stimulate growth and improvement in teaching through reflective practice but evaluation is used for summative purposes to give an analysis of a particular teachers' method in a particular course and setting.

2.3 Types of Test and Testing.

According to Nitko, testing is a systematic process to observe persons and describe their progress with a numerical scale or a category system. Thus, tests may give either qualitative or quantitative information. A test or examination is an assessment intended to measure test-takers knowledge or skill in many other subjects. It can be administered orally, written, on a computer that requires a test taker physically perform a set of skills. Moreover, testing is a system, which practice to make an objective judgment that shows passing or failing to meet stated objectives. In this sense, there are two crucial purposes of testing, which are about affirmation that, what was specified is what was delivered: it refers to the product (system) that meets the functional, performance, design, and implementation requirements identified in the acquisition process. Second, testing is a kind of taking risk for both acquiring the activity and system's developer. Therefore, tests should be included into curriculum at schools; they are inevitable elements of learning process. It shows both the learners and the language teachers how much the learners have learnt during the learning process. To check the learner's intelligence the teacher can apply a lot of techniques. Heaton (1990) states that, tests can be used to show the positive and negative effect of teaching style or method and help the teacher to make it better. Moreover, test results will show the learners weaknesses, and if it is carefully managed by the teacher, the students will make progress. Tests should be referred to a set of items or questions under specific conditions. Researchers both believe students learn more when they have tests and have some difficulties while testing. Certainly, while learners preparing for the tests they have to practice all the

materials that are supposed to be tested. However, it is not the matter of acquisition and productivity in this type of learning. On the contrary, it often causes anxiety and result in a stressful situation that the learner found him/herself before or during the test.

On the other hand, the tests can sometimes promote the students' acquisition process. Constantly being able to test by small tests, students can learn the subject properly and transfer it to his/her long-term memory. However, according to Thompson tests reduce practice and teaching time. This causes limitation and lack of practice. Hence, it is definitely recommended that syllabus should be observed so learners find themselves in precise frames that the teacher will employ. Whereas, tests can offer advantages like increasing learning, the students are supposed to study harder during the preparation time before a test. As Alderson (1996, p:212) figures out, educators should not forget that during administrated tests students get little support from the teacher compared to the support they get in classroom activities. The students have to manage themselves; they cannot depend on the help of the teacher. While students are expected to do the exercises with the related materials, they can share their ideas with their partners or they become aware they can get the teacher's help if they require it, but while assessment that is inapplicable. The researcher believes that the teacher should teach the students to overcome their fear of tests and assist them to obtain the ability to work by one self-related to their own responsibilities and knowledge. In this sense, success triggers learners to study more, encourages them to go ahead of even rather difficult tests. Therefore, it is clear that tests can be seen as a tool to increase motivation and self-confidence. Although, students try reasonable times, the student would definitely lose their confidence and eagerness.

Hicks (2000) states that especially in language learning the role of test and evaluation is very valuable and beneficial. Testing does not refer to the teacher's desire to catch the students when they are not ready for the assessment; it is also not the encouraging factor for the learners to study. Literally, the test is a requirement for information and a chance to learn what the teachers want to know about their students' background knowledge. On the other hand, the test is supposed to demonstrate the students' weak points, and their strong sides. Moreover, according to the idea of Hughes (1989, p.5) language teachers can also check the development of

their students and obtain general or specific knowledge about them. This notion straightforwardly leads us to comprehend that for each of these purposes there is a special type of testing.

In addition to all those features of testing and assessments some scientists (Hughes, 1989; Heaton, 1990; Alderson, 1996; Underhill, 1991; Thompson, 2001) declared that there are four common types of tests in accordance with their usage: proficiency tests, achievement tests, diagnostic tests, and placement tests.

Hughes (1989) states that *Proficiency Tests* are designed to measure people's ability in a language; regardless of any training, they may have had in that language. This type of test aims to measure test takers' readiness for a particular communicative role and tries to analyze how much of the subject the students enroll or acquire. There is no mandatory curriculum or syllabus for this kind of test but it is intended to check learner's language competence. Rather than giving more importance to some preparation and administration issues, the results of the test are what being focused on. The Test of English as Foreign Language (TOEFL) can be given as an example of proficiency test. This exam measures the students' general knowledge of language in order to allow them to introduce themselves any high educational institutions or to get an occupation in the United States. Cambridge First Certificate (CFC) is another proficiency test that has the similar goal as the previous exam type. Hughes (1989) emphasizes the similar description of proficiency tests focusing that education is not the thing that is asserted, but the language. He also stated that 'proficient' in process of proficiency tests stands for preparing a certain ability of using the language according to the relevant purpose. In this sense the researcher expresses that the learners' ability could be measured in diverse fields or subjects in order to check if the students could meet the requirements of a specific area or not. According to Hughes (1989), the only similar factors about such tests are not depended on any courses, but they are all intended to measure the learners' appropriateness for definite course at the university. Since there were a lot of items to overcome both language teachers and learners may have difficulties to handle this issue. All four skills are being practiced during the preparation course by using different methods with possible activities.

An *Achievement test* is another type of test that is directly related to the courses and their initial purpose is being to determine how successful individual

students, groups of students, or the courses themselves have been in achieving objectives. Achievement tests measure a language if someone has acquired the information throughout the course, or study. Alderson (1996, p.219) indicates that achievement tests are “more formal”, whereas Hughes (1989, p. 8) declares that this type of tests, needs teachers involvement for the preparations. The reason of most language programs is to increase the level of language perception and observe their development; therefore, most of the language teachers find themselves interested in making achievement tests to assess their learners’ ability. In this sense those tests involve who will be advanced to the next level of study or which students should graduate. Achievement tests help language teachers to make realistic decisions that will help to improve achievement in their language programs or they may find a need to make and confirm changes materials, equipment, curriculum design, and so on. This type of tests can be design with very specific reference to a particular course. Therefore, this specific program means that the achievement test will be directly based on course objectives and for this reason it is criterion-referenced. Such test may be beneficial and administered at the end of a course to decide how effectively learners have mastered the instructional objectives. Achievement tests not only be used to make decisions about students’ level of learning but also be used to affect curriculum changes and to test those changes.

One other type of test is a *diagnostic test*, which is used to identify learners’ strengths and weaknesses. Language teachers implement diagnostic tests when they want to know how much learners know about the subject so that teachers have chance to analyze what to do next. Diagnostic tests can be given at any stage during a course to help teachers plan future lessons. This type generally consists of several short but reliable subtests, which measures different language skills of a single broad skill. According to Alderson (2005), “Diagnostic tests are based on some theory of language development, preferably a detailed theory rather a global theory”.

The core of diagnostic testing lies in a comprehensive and creative feedback system as in most of the assessment types. In this regard, Alderson (2005) asserted ‘The provision of feedback to students after they take a test is one of the unique features of diagnostic tests.’ Diagnostic feedback can be also helpful for the learners in understanding what objectives have not been achieved. It also enables to set specific goals to improve their language autonomy. On the other side, feedback can

assist language teachers better reach teaching objectives and make compatible instructional adjustments. Both proficiency tests and placement test can serve for diagnostic objectives and it is very difficult to distinguish diagnostic test from other kinds of assessment. In order to distinguish diagnostic test from other types of tests, Alderson (2005) asserted the characteristic of the diagnostic testing through its features. Those features of diagnostic tests include the major aspects of language test development like; test purpose, test construct, test content, test feedback, etc. Moreover, this may encourage language testers to make further efforts to improve the design of diagnostic tests. Hughes also presumes that diagnostic tests are difficult to prepare. Therefore, very few tests are constructed just for diagnostic purposes. As Hughes (2003) suggested, this type of test is advantageous for individualized instructions. It means that it is useful for checking a certain item; it is not necessary to cover comprehensive issues of the language.

Placement tests are necessary to provide information, which helps to place learners at the stage of the language learning programme most appropriate to their abilities. When the learners first arrive at the language institution, teachers give students placement tests in order to understand what level they should study at. According to Brown (2007, p.454) ‘placement tests typically includes an example of material that includes in the curriculum, and it provides an explanation of the point at which the student will find a level or class to be neither too easy nor too difficult, but appropriately challenging. This may also help to put the student exactly in a group that responds his/her true abilities. Heaton (1990) stated that type of testing should be general and should purely focus on a vast range of topics of the language not on just specific one. The placement test should deal exactly with the language skills relevant to those that will be taught during a particular course. The importance of the syllabus was indicated and should be analyzed beforehand. Hughes (2003) emphasizes that all foundations have their own placement exams according to their learner’s requirements. In order to put the students into different groups, which are relevant to their level, language teachers need the learners’ exam results.

2.4 Types of Assessment Summative &Formative

Assessment for learning can contribute to the development of effective schools. If assessments *of learning* provide evidence of achievement for public

reporting, then assessments *for learning* serve to help students learn more. (Stiggins, R. J. 2002) Assessment themselves have been despised. The most important issue is why assessments are given and how the data is used. Assessment is the measurement of what students are learning. It can also be defined as how well they have mastered certain target skills. On the other hand, assessments provide language teachers with both subjective and objective information in order to confirm student progress and development of their skills.

Assessments can be either summative or formative and it is very significant to make that clarification. As Hughes stated (2003), assessment is formative when teachers use it to check on the progress of their learners, to see how far they have mastered what they should have learned, and then use this information to modify their future teaching plans. The major goal of formative assessment is to observe students learning and to provide continuous feedback that can be used for the language teachers both to improve their teaching style and for the students to improve their knowledge and perception. Therefore, formative assessment is appropriate to help students analyze their strengths and weaknesses. This type of assessment also helps institutions to recognize where students are tackling and address problems directly. Formative assessments have low or no point value it gives information both teachers and students about student understanding at a point when timely arrangement can be made. The arrangements help to guarantee students achievement, aimed at standards-based learning goals in a set time frame. Rather than measuring the students' grades with four skills, this assessment type assist teachers determine next steps during the learning process.

As Hughes (2003) stated summative assessment should be applied at the end of the quarter or semester to have a look at the progress. These assessments are given to decide at a particular point in time what learners know and do not know and it was given periodically. End-of-unit or chapter tests, semester exams or end-of-term exams can be given as an example for the summative assessment. The goal of summative assessment is to evaluate students learning process and compare it against some standard or reference point. However, the information, which is gathered from this type of assessment, is important; some researchers claimed that it can only help in evaluating certain aspects of the learning process. Since they appear after instruction every week, months, or once a year, summative assessments are tools to

evaluate the effectiveness of programs, school improvement goals, schools' curriculum. According to Alderson (2005), summative assessment is a long traditional test, which was so stressful to students. Any kinds of test, which is only possible use of gathering scores in the eyes of students, can be summative assessment although teachers have primarily designed the test to promote learning and teaching.

Formative and summative evaluation is the important elements. Although these two types of evaluation, functions, uses and purpose differ from each other, it cannot be said that one is superior to the other and both of them are necessary for educational programs to be effective. This current study quite fits the frame of formative assessment because it tries to clarify whether the program is working well or not, and which improvement is needed.

2.5 The Effects of Assessment and Evaluation on Learning

Evaluation and assessment has a crucial impact in the journey to success. It also affects students learning process, which is interested in getting information about learners' perception, then analyze concerning their achievement. Testing and evaluation can be one of the significant forms of getting information on the way of teaching and learning. Any of the feedback or information acquired is the reaction of learning. Students' learning outcomes not only demonstrate students' success but also reflect the success of institution in managing the learning practice. The aim of the evaluation should be getting accurate information about the level of the learners, which is connected, to students learning objectives. Evaluation has a link between learning to provide students information and guidance in order to assist them to plan and manage their next steps in their learning. On the other hand, assessment also helps language teachers to adjust their teaching strategies, and identify students' learning needs in a clear and practical way. When language teachers determine what their students learn, they should also need to determine the way of evaluation at the end as well as assessing their development. In higher education, language teachers should care about the students' assessment and feedback. Learners are often assessed to see how well they have done or how well they perform. As Brown (2005) stated 'Nothing we do to, or for our students is more important than our assessment of their work and the feedback we give them on it. The results of our assessment influence

students for the rest of their lives and careers'. Assessments are significant for the learners to provide feedback, to promote their learning, and to analyze their strengths and weaknesses. Evaluation and assessment should integrate individual differences in students. In this sense, a wide range of assessments should be implemented, in order not to damage any individual or group of learners.

Assessment for learning focuses on opportunities to improve learners' ability to make judgments about their own performance and try to make better upon it. Through both formative and summative assessment methods, students develop their skills. Well-designed assessment has lots of benefits to engage students with their learning. According to Race & Brown and Smith's study (2005), assessment can strengthen active learning especially when the assessment delivery is creative and charming. The assessments provide number of skills such as developing critical thinking, self-awareness and reflection. Therefore, diverse assessment methods can be helpful for both learners and language teachers to analyze their language awareness. Ramsden (2003) claimed that using a range of assessment methods gives students more latitude to demonstrate their knowledge and skills across a range of contexts. By implementing wide range of assessment, learners have a chance to notice the missing parts and lack of their skills while learning and testing their language competency.

2.6 Teachers' and Students' Perceptions of the Exams

According to Ecclestone and Pryor (2003), the idea of learners' assessment careers was a way of understanding the impact of different assessment systems on learners' dispositions in various learning contexts. Researchers claimed that students' personal identities and learning are shaped by consistent and powerful assessment leadership of organization. When it is considered the characteristics and impact of assessment from the students' points of view, learners' perceptions about testing and assessment have significantly influences on students learning. Their perceptions can be changed depending on the type of exams and questions. Brown and Wang (2013) claimed that 'the beliefs, attitudes, understanding and practices students develop around assessment through their experience of an assessment regime constitute their assessment career. In this sense, the attitudes towards the assessment for the learners depend on learners' experiences through the learning process. To believe assessment

in order to improve language perception approves a learner to self-directed learning. In their evaluation, learners use errors and mistakes as a mechanism to improve language learning. It is important to ensure that personal attitudes and behaviors, goals and intentions influenced by and demonstrate the aim of learners towards the assessments. The research, which was conducted, by Struyven, Dochy, and Janssens (2005) assisted, in their review of the research literature on student perceptions of assessment in higher education, that how a learner perceives an assessment will influence how that student learns and studies. For instance, in their studies the researchers found that the need to avoid test-anxiety resulted in a preference for multiple-choice questions. On the other hand, other studies related to students perceptions in European universities found that students' learning-strategy preference help them to shape their understanding of the learning required by assessments. Zilberberg et al. (2009) points out that 'unless students understood the purpose of assessment, their test-taking motivation would be greatly depressed. Series of studies conducted in New Zealand through the Student Conceptions of Assessment (SCoA) and analyzed the significant conception of assessment (Brown 2008; Brown & Hirschfeld 2007, 2008; Brown et al. 2009; Brown, Peterson and Irving, 2009). In reference to the analyses, assessment leads to improved teaching and learning; assessment relates to external factors such as school quality and student futures; assessment has a positive emotional and social impact on students; on the other hand, the result of the study also shows assessment is irrelevant because it is bad and can be ignored. Generally, learners admitted that assessment was for 'improvement' and refused its irrelevance. Some research studies (Pekrun, Elliot, & Marier 2006; Pekrun, 2002), which was implemented for university student achievement emotions have analyzed that learners have more emotions than anxiety. There were positive (e.g. hope, enjoyment, pride) and negative (e.g. anxiety, shame, sadness or boredom) emotions from the responses. As it was stated in the result of the study (Pekrun et. al.2006), while the qualifications of the schools increase, students' dominant emotional response to assessment appear to become increasingly negative.

Cheng (2008) states the significance of high-stakes examinations, which has an extremely effective impact on learners, and huge impact on what is to be taught and learned. Test of English for Academic Purpose developers declared their

assumption that new type of tests draw attention to productive and cognitive skills, which leads better language learning. On the other hand, examinations in Korea were designed ‘to activate the teaching of speaking and writing of English at schools, which has not actually been conducted despite its compelling necessity’ (Lee, 2012, p.30). As can be seen clearly language tests are used to measure abilities and skills that test developers focus on students learning activities and developments in their learning environment.

From the perception of teachers, assessment and evaluation plays a crucial role in the process of teaching and learning. Moreover, they provide the significant part of curriculum and determine the content of learning. The right evaluation or assessment can provide beneficial information for students, teachers, parents, and schools. In this sense, those information from assessments may help language teachers decide which instructional methods are best for certain students, what their learners may already know about the given subject, and what subjects need to be taught again. According to Palmer and Bachman (2010), language testers need to take into account the intended consequences of the test in test-design. In this context, test users’ perceptions of the evaluation or abilities measured in the test are examined to play a significant role in accomplishing the *washback*. Considering the current study the crucial role that teachers play in the assessment process light the way for the causes and effects of implementation in educational context. Rea-Dickins and Gardner (2000) conducted a survey about teachers’ ideas towards formative assessment through interviews. The study was implemented to find out that teachers benefited through providing evidence regarding students’ learning; planning and managing their teaching; identifying the development extend for teachers and students alike as determined by curriculum; and providing feedback on their own teaching(Rea-Dickins&Gardner, 2000).

In the study carried by Troudi et al., (2009) conducted to examine the language assessment and teachers’ own role in the implementation of second language assessment in the United Arab Emirates and Kuwait. The result of the study indicated that EFL teachers’ conceptualizations of the role of assessment as well as their own role in assessment are based on their knowledge of the field, the contextual environment, and employment policies. Accordingly, teachers’ education and experience, and their own personal beliefs and values affect their language

assessment. As it is also known, assessing learners' performance is one of the most critical aspects of the language teachers, however many teachers in the United States do not feel adequately prepared to assess their students' performance (Mertler, 1998, 1999; Stiggins, 1999). Studies have shown that, the result of feeling of discomfort and inadequate preparation by teachers' is the reason of limited assessment literacy (Popham, 2003). While designing assessment tools it should be considered reasonable both for students' preferences and teachers' justification might influence the students' learning and the way they tested. According to Plake (1993), teachers often claim that their lack of preparation is largely due to inadequate pre-service training in educational measurement. In the light of this study, Mertler (1999) conducted a research through asking in-service teachers about their perceived level of preparedness to assess student learning resulting specifically from their teacher preparation programs. Reported responses have shown that they were not well prepared.

To sum up, students and teachers' perceptions enrich the language assessment in many respects. The perceived methods of assessment seem to have a significant impact both for teachers and students' approaches. The current study aims to investigate assessment content from the learners' perspective.

2.7 Previous Studies on Language Assessment and Students' Perception

There are various studies aimed to evaluate the effects of assessment from the teachers and the learners' perspectives in EFL contexts. Although sufficient number of studies highlighted on testing and evaluation in language teaching but there has been little experimental research done on students and teachers' perception in language assessment.

Alderson and Wall (1993) aimed to demonstrate the effect of exam that is brought in Sri Lanka on language lessons. Data were collected thorough the observations. During the observations, the researchers realized that language teachers spent more time on the skills, which were tested in the exams. The findings of the study demonstrate that researchers noticed negative washback due to the teachers' unequal attitudes towards the content of the test.

To begin with, Mussaway (2009) conducted a study to explore faculty members' perceptions of classroom assessment and their expectations of students' learning. The researcher was also concerned about learning and the extent to which assessments results were used to improve students' learning and classroom instruction. Two types of data were collected through the interviews with focus group of students and teachers, 16 hours observation, and ranging scaling questionnaire provided 209 participants. According to the results, both students and teachers had recognition of various forms and purposes of classroom assessment. Responses from the students' perceptions demonstrate that they have the same viewpoints regarding the definition and the purpose of classroom assessment.

In other similar study carried out by Pekkanlı (2010), 'Traditional Foreign Language Classroom Assessment in Turkish High Schools' with the aim of investigating the current situation of language testing through test development and grading procedures. The data were collected from a questionnaire and interview based on language testing. The findings indicated that the current procedures of foreign language testing in Turkish high schools are problematic in terms of test content, test tasks, language skills tested, fairness of tests, reliability and validity. In her study, she recommended teachers to be conscious of the importance and consequences of testing.

Duran (2011) carried out a study to explore language instructors and learners' perceptions of the washback effect of a speaking test in a classroom atmosphere. The researcher conducted the study with 307 intermediate level and 45 instructors of English through questionnaires, and interviews. The findings explained the positive manners of teachers and students' concerning the significance of teaching and assessing speaking. Considering the washback affect, not only teachers but also learners believe that they would continue to teach and learn speaking skills in class even if there was no assessment on speaking skills.

Momeni and Barimani (2012) examined in their studies whether testing frequency is beneficial or harmful on language achievement of Iranian EFL learners at pre-intermediate level. Pre-tests and post-tests were adopted to gather the data. The research has two essential results. First, frequent testing has a positive washback effect on language achievement. Second, it made students eager to learn.

In brief, most of the studies in the literature have examined thoroughly the assessment system. Whereas, there is always need to comprehend the effects of testing on language learners. This present study aims to fill the gaps by exploring EFL students' perceptions of the content of tests and frequent testing.

Chapter 3: Methodology

3.1 Overview

This chapter aims to describe the methodology of the study by giving information about the research questions, philosophical paradigm, research design, setting, participants, data collection instruments and procedures, data analysis, and limitations of the study.

This study attempts to find answers to the following questions,

1. What are the general perceptions of test content by students?
2. What are the general perceptions of tests by different levels of students?
3. How do students perceive the testing practices in terms of frequency, compatibility, clarity, length, complexity, variety?
4. What are the students' perceptions of frequent testing?
5. What are the students' perceptions of consistency of testing with teaching practices?

3.1.1 Research context. The assessment system at Preparatory School is diverse. Learners start the academic year with the placement test, which is administered by the Preparatory School at the very beginning of the term. Students attend four models in one academic year that continuous (7) weeks. In order to be successful at the end of the quarter learners need minimum GPA of 70/100. The percentages of the assessments are as follows:

Table1

Module Assessment Template (2015-2016). Taken from Preparatory School students' handbook

2015-2016 Academic Year	
8%	Participation
8%	Reading Circle
8%	Weekly Writing Assessment
8%	Weekly Online Speaking Assessment
8%	Weekly Listening Assessment
20%	2 Achievement Exams
40%	Exit Examination

Table 1 shows the list of the module assessments in a whole year at Preparatory School. All new learners must verify that their English proficiency levels are adequate to follow standard courses followed in the university. At the very beginning of each academic year, learners take the Level Placement Test. Students who exceed the minimum required scores can enter the Proficiency Exam. Those who get the minimum required scores are excused from having attended the Preparatory Class. On the other hand, students who do not accomplish the minimum scores are required to attend the Preparatory Class. For each quarters students assessed by the followings above which is shown in table 1. Learners attend four (4) modules in one academic year. Each module lasts seven (7) weeks. To complete the quarter successfully, students need to accomplish a quarter with a minimum GPA of 70/100 in order to continue to the next level.

Placement Test

All students entering the Foreign Language Preparatory Class must take the Placement Test. The Placement Test is made up of 100 multiple choice questions. The test measures students' reading and use of English skills as well as their knowledge of vocabulary.

Proficiency Exam

Students who get or exceed the required score of 70 on the Placement Test may sit the Proficiency Exam. Students who obtain a passing score (70) in this exam may move on to their respective departments. The Proficiency Exam is made up of four sections, which are, listening (20%), reading (35%), writing (25%), and speaking (20%).

In order to pass the Preparatory Class learners have two alternatives. The first option is, students need to successfully complete the B1+ (Intermediate Level 2) module which means learners should get a minimum final score of 70 out of a 100 for the module. Another option is to take the Proficiency Exam twice a year - once at the end of the Orientation Programme and once before the start of the third quarter. Students who meet or exceed the required score for the Proficiency Exam may continue their education in their departments.

Exit Exam

These are written assessments, which are taken at the end of the quarter. Exams include all the information taught throughout the seven (7) weeks. The result of the exams show whether a student is at the level for the following module.

Make-up examinations

Make-up examinations are given to students who have missed an exam. Only if the students submit a medical report for the day of the exam they missed can take the make-up examination.

3.2 Philosophical Paradigm

Mertens (2015) declares, “Paradigm is a means of looking at the world. It is consist of certain philosophical assumptions that guide and direct thinking and action’. Cohen and Manion (1994, p.38) describe paradigm as ‘the philosophical intent or motivation for undertaking a study.”According to Mac Naughton, Rolfe and Siraj-Blatchford (2001) provide a definition of paradigm, which includes three elements: a belief about the nature of knowledge, a methodology and criteria for validity. Creswell (1994) also argues that, quantitative studies base their analyses on measurements, numbers, tests, while qualitative research has a complex and comprehensive approach that was asserted through words and reports of the people in their point of view. While the major paradigms will have an overall framework consistent with the definitions provided above, specific research paradigms might have particular features, which differentiate them from other paradigms within the same group. For educational researchers, several major paradigms govern their analyses into the policies and practices of education. Each paradigm carries related theories of teaching and learning, curriculum and assessment, professional development, etc.

In this study, the combination of qualitative and quantitative methods, which refers to mixed method paradigm, was applied for pragmatic reasons and promote both methods’ strengths and weaknesses. Pragmatism is not committed to one system of method or philosophy. Pragmatist researchers focus on the 'what' and 'how' of the research problem (Creswell, 2003, p.11). While pragmatism is seen as the paradigm that provides the underlying philosophical framework for mixed-methods research (Tashakkori&Teddlie, 2003; Somekh&Lewin, 2005), some mixed-methods researchers align themselves philosophically with the transformative paradigm (Mertens, 2005). It this sense, it is clear that mixed methods could be used with any paradigm.

3.3 Research Design

At Preparatory School, there are a lot of evaluations and assessments for four skills. Some of the exams like achievement, placement, listening assessments are prepared by the testing unit and the coordinators. Yet, no satisfactory study or

measurement has been made to comprehend the effects of such testing or what type of assessment can be used to help students' learning. This study aims to find out if the learners' perceptions of the exams are working the way it has been arranged, getting feedback from language learners for the exam types and the improvement. In this case study, mixed method has been conducted. For the quantitative part of the data questionnaire was implemented in two sections, while for the qualitative data interview was done with both the learners and the teachers. The analysis of the data was based on both qualitative and quantitative method to evaluate the testing instruments at Preparatory School in Izmir, Turkey.

3.4 Setting

The current research study is carried out to evaluate the effect of testing and assessment on language learning at Preparatory School of Gediz in Izmir, Turkey. The Preparatory School is designed to assist students in their journey of the English language. Foreign Language Preparatory School program delivered in 7-week terms. In Preparatory School courses, students learn to develop the four skills as well as communicative skills to use during their academic studies. Learners have three different instructors for every module. In this way, students are exposed to a variety of teaching styles, accents and personalities.

Through inductive teaching methods and chances for collaborative and interactive learning, the program aims to provide students with the necessary skills required not only for their departments, but also for their future careers. The program consists of a modular system made up of four modules; A1 (elementary), A2 (pre-intermediate), B1 (intermediate), and B1+ (intermediate). Although there is no separate lessons dedicated entirely to four skills, throughout each of the levels, students are engaged in a variety of tasks, assignments and activities focusing on their oral production skills, writing skills, listening and reading skills. There are pop quizzes for listening skills, weekly participation grades that were given by language teachers to assess the quality and quantity of language production in the classroom, completion of homework assignments, and active classroom participation during lessons. There are also weekly writing assessments, online speaking assessments, two achievement exams and exit examination in Preparatory School. As students'

progress from the beginner level to the upper-intermediate level, the level groups are formed according to students' exit examination.

3.5 Universe and Participants

In the present study, data were collected from 124 participants who currently studied at Preparatory School, in Izmir. The participants were chosen through quantitative methods and for each level (A1 beginner, A2 elementary, B1 intermediate, B1 + intermediate), the whole class members were chosen randomly. One group of the participants (A1 repeat) had proven unsuccessful in moving onto the next level of their education, and therefore their language development begged for more understanding. The propellant to decide this participant group was based on the observations that the language teachers shared with the researcher. The larger group from which participants were chosen, a quantitative technique was used in order to identify the participants for the study by random sampling. Their age range was 18-22 and they have diverse educational experience about language learning.

3.6 Procedures

In this section types of sampling, data collection tools, data analysis procedures, trustworthiness, and limitations were provided in detailed.

3.6.1 Source of Data. Sampling is the method of selecting the number of individuals for a study in such a way that the individuals represent the larger group from which were selected. The purpose of the sampling is to gather data about the population in order to make an inference that can be generalized to the population. Creswell noted that in qualitative research, "the intent is not to generalize to a population, but to develop an in-depth exploration of a central phenomenon", which is best achieved by using purposeful sampling strategies (2005, p.203). Identifying the population, defining the sample size, selecting the samples are the crucial issues of quantitative sampling. There are two types of sampling techniques, which can be categorized in probability, and nonprobability sampling. While probability sampling uses random selection, non-probability sampling does not involve random selection and methods are not based on the rationale of probability theory.

Qualitative and quantitative research methods are different from the aims of their methods and sampling techniques as well. In the present study, random sampling was used, which allows individuals an equal chance of being take part in the samples. As Creswell stated, ‘the most popular and rigorous form of probability sampling from a population is simple random sampling.’ (Creswell, 2014, p.161). In simple random sampling, all participants have an equal and independent chance of being selected. In this sense, it is easier to conduct the study, and this meets the assumptions of many statistical procedures.

3.6.2 Data Collection Procedure. The study aims to triangulate data collection by presenting various outlets through which data was obtained. The data for this research were obtained through three sources: Questionnaire comprised the quantitative part of the research on the other hand, interview and document analyses constituted the qualitative data.

3.6.2.1 Questionnaire. In this study, the questionnaire was conducted within the aim to see how effective test and assessments in language learning from the learners perspective. This questionnaire was prepared to get large amounts of information from a large number of people in a short period of time and in a relatively effective way. The questions were prepared to examine what extents do each of the exams help learners to comprehend language in Preparatory School of Gediz, in Izmir. The questions were related to each part of the exams including listening, grammar, reading, vocabulary, and writing. The reliability of the students’ perceptions of testing and assessment questionnaire was evaluated. The data show that questionnaire has high reliability 0.915 in the 24-item.

Table 2

Reliability Statistics

Cronbach’s Alpha	N of Items
,915	24

3.6.2.2 Interview. In order to provide useful information when researcher cannot directly observe participants, having an interview is the appropriate way to describe detailed personal information. Interviewing has many different types such as individual, face-to-face group or individual interviews and written interviewing.

In this study, 4 open-ended questions were asked to get data from the interview. The interview was done with the students and language teachers to permit them to share their experiences and attitudes about types of questions, language teachers and learners' feedback related to exams and effects of exams on students' language education.

Table 3

Overview of Research Questions and Corresponding Procedures

Research Question	Data Collection Instruments	Data Analysis
1. What are the general perceptions of test content by students?	Questionnaire	SPSS (Cross tabulation)
2. What are the general perceptions of tests by different levels of students?	Questionnaire	SPSS (Cross tabulation)
3. How the students perceive the testing practices in terms of;	Questionnaire	SPSS (T-test Anova)
a. Frequency		
b. Compatibility		
c. Clarity		
d. Length		
e. Complexity		
f. Variety		
4. What are the students' perceptions of frequent testing?	Interview	Content Analysis
5. What are the students' perceptions of consistency of testing with teaching practices?	Interview	Content Analysis

3.6.3 Data analysis procedure. In this study, both qualitative and quantitative data were gathered and analyzed. The questionnaire was implemented at the very beginning of the study in order to get information about participant general attitudes toward the research topic. In the first section of the questionnaire, open-ended questions were used. Participants marked their answers to specific open-ended questions. All questions were prepared in order to elicit information in the same direction. In the second part of the questionnaire, an open-ended questionnaire was used to collect extensive data from participants in comparison to using a Likert scale. The researcher developed a questionnaire based on the needs of the current study, which aims to investigate testing effects on language learning. The aim of the questionnaire was to obtain data in two significant issues relevant to the questions of this study. The first issue was the effect of exams in language learning, and the second issue was the other factors, which are compatibility, frequency, variety, length, complexity, and clarity of the exams that affect the needs of learners. Data were analyzed by the Statistical Program for Social Science (SPSS). For the first section of the questionnaire the data were interpreted by cross tabulation to get the general perceptions of test content, while for the second section of the questionnaire the data were analyzed by doing T- test to get the perceptions of learners' for the specific contents of tests.

In qualitative part of the research, interviews were analyzed. The focused group interview was used to collect shared understanding from several individuals as well as to get impressions from specific people. The focus group consisted of five students and for each level. The researcher asked a question related to the coherence between the content of the exam items and classroom practices. Open –coding was implemented through data analysis done by hand as well as line by line for every sentence, even word by word. In this sense the researcher give voice and meaning around an assessment context.

3.6.4 Trustworthiness. Lincoln and Guba (1994) suppose that trustworthiness of a research study is essential to evaluate its worth. Researchers seek to satisfy four criteria, which is addressing credibility, transferability, dependability, and conformability. The first provision credibility is a kind of confidence in the truth

of findings. Credibility depends more on the richness of the information gathered, rather than the amount of data. There are many techniques to measure the accuracy of findings such as member checking, triangulation, case analyses, peer debriefing...etc. In this current study triangulation was used to increase the research credibility, which refers to multiple data source. One of the data sources of this study was questionnaire to examine the effect of assessment system on language learning. The questionnaire, which aims to search the content and the effects of assessment on learning, was prepared by the researcher. In order to support and assess the trustworthiness the questionnaire of the current study was evaluated by the Cronbach's Alpha statistics. Another data collection source of the study was interview. Four open-ended questions were asked both the language teachers and learners. The data were analyzed by the content analysis method. By using three different data collection method, the current research had a multiple perspective and it provides different aspects of the research focus.

Transferability means generalization of the research findings to other situations and context. Transferability does not include broad claims, but invites readers of research to make connections between elements of a study. This study put emphasis on the effects of assessment on language learning for students' perspective in detailed. Therefore, the current study can be a model for the future testing and assessment studies from the learners' perspective only.

Dependability is the processes of the study that should be reported in detail, through enabling a future researcher to repeat the work, if not necessarily to gain the same results. If the work were repeated, in the same context, with the same methods and with the same participants, similar results would be obtained and it means the study is dependable. To develop a thorough understanding, the study includes comprehensive and interrelated data collection and data analysis part to ensure the consistency of the results.

Conformability is the last essential part of trustworthiness, which is a degree of neutrality or the extent to which the findings of a study are shaped by the respondents and not researcher bias, motivation, or interest. To address this, the researcher conducted member-checking, which involved one of the researcher colleagues to confirm the themes that emerged and debriefing which involved the participants' confirmation of what they reported in the written interviews.

3.6.5 Limitations. One of the major limitations in this study is time management for the implementation of the questionnaire. Filling out a research questionnaire was done a few days before the achievement exam; therefore, fewer students were presented at that time to get the perceptions. Motivational level can also be a factor. The reliability of findings could be improved by including more participants in the study.

Chapter 4:

Results

4.1 Overview

This chapter focuses on the respondents' answers in terms of the effects of exams on students' language learning in school of foreign language in Izmir, Turkey. In this regard, the current study tries to find out the effectiveness of evaluation and measurement from both the students' and teachers' perspective and provide some relevant and reliable suggestions. Data were collected through the questionnaire and interview, which constitute qualitative and quantitative results of this chapter. The following section explains the findings related to each research question existing in the study.

4.2 Quantitative Section -Results

4.2.1 The findings of the general perceptions of test content by students.

In view of the first research question, the data were obtained through the questionnaire. Initially, the students were asked to evaluate general perceptions of test in terms of five skills such as; vocabulary, reading, listening, writing, and grammar, general perceptions of test by different levels, and the perception of testing practices in terms of consistency, frequency, diversity, length, complexity and clarity. In order to get the general perceptions of the students on test content the following questions were asked for the first part of the questionnaire;

Q1. What is your favorite part in the exams?

Q2. Which part of the exam do you think is the most helpful and has the biggest effect on your final result?

Q3. From which part do you generally take the highest point in exams?

Q4. Which part of the exam is the most difficult?

Q5. From which part do you take the lowest points in the exams?

Q6. What is the most time consuming part in the exams?

Table 4

The General Perceptions of Test Content by Students

	Vocab.	Reading	Listening	Writing	Grammar
Q1.	34.0	19.8	16.3	17.0	12.7
Q2.	29.4	14.3	17.1	30.1	8.9
Q3.	20.0	24.8	15.1	26.2	13.7
Q4.	8.57	14.2	35.0	22.1	20.0
Q5.	17.6	13.0	34.6	10.7	23.8
Q6.	2.1	44.6	3.5	46.0	3.5

Table 4 reflects the perceptions of test content in five skills. Based on the obtained percentages, 34% of the students like the vocabulary part in the exams, while almost half of this percentage (12.7%) is for the grammar parts. When the activities which are done in the classroom compared with grades vocabulary and writing had almost the same percentages, (29.4% and 30.1%), while the least helpful part for the grades and effective part of the scores was chosen by the grammar section which is 8.9%. As indicated in the table above, the participants got the highest mark from writing parts (26.2%), but they got the lowest mark from the grammar part (13.7%). For the other skills, there has been an approximate distribution. The majority of the participants found the listening part most difficult (35%), while 8,57% of them marked vocabulary part the most difficult one. When the participants were asked which part of the exam, they got the lowest grade, among the other skills listening part was chosen, at (34.6%). Although, almost half of the participant (46%) reflected writing section took most of their time in exams, 2.1% participants marked vocabulary as the most time consuming part in the exams.

Table 5

The Differences between the Levels of Students' General Perceptions of Tests.

Considering the second research question, the following analysis aims to investigate the content of the assessment and evaluation by different levels of students.

		Vocabulary	Reading	Listening	Writing	Grammar
Q1.	L	28.57	17.46	12.69	26.98	14.28
	H	38.46	21.79	19.23	8.97	11.53
Q2.	L	29.68	15.62	12.5	28.12	14.06
	H	29.26	13.41	20.73	31.70	4.87
Q3.	L	11.94	31.34	11.94	26.86	17.91
	H	26.92	19.23	17.94	25.64	10.25
Q4.	L	9.67	6.45	43.54	29.03	11.29
	H	7.69	20.51	28.20	16.6	26.92
Q5.	L	19.64	5.35	44.64	16.07	14.28
	H	16.21	18.91	27.02	6.75	31.08
Q6.	L	3.12	32.81	6.25	51.56	6.25
	H	1.3	54.6	1.3	41.3	1.3

For the first item both high levels and low levels have the same response, for vocabulary, which is at 28.57% for low, and 38.46% for high levels. However, participants marked listening (12.69%) for low levels and writing (8.97%) for high levels as for the low percentages. This means that while they like to do vocabulary questions in the exams, they have some concerns about listening and writing. As for the second item, 29.68% of participants marked vocabulary for low levels, 31.70% of participants marked writing for high levels as the most effective and helpful sections for their grades. On the other hand, students marked grammar section at 4.87% for the high levels and listening section for the low levels as a lower percentage. The following item aims to investigate the participant's highest grade part. Reading part (31.34%) for the low levels and vocabulary part for the high level

marked as a response of this item. As for the fourth item, both high levels at 43.54% and low levels at 28.20% marked listening section for the most difficult part in the exams. The subsequent item, the students were asked which part of the exam they get the lowest mark. Almost half of the low level participants (44.64%) marked as listening while 31.08% participants marked as grammar. The last section aims to investigate the most time consuming part of the exam. More than half of the participant both in low level (51.56%) and in high level (54.6%) marked but variously low levels marked writing section and high levels marked reading section as the most time consuming part.

4.2.2 The findings of differences between the levels of students' general perceptions of testing and assessment. As for the third research question regarding to evaluate the students' perceptions of testing practices in terms of compatibility with the teaching practices (F1), frequency of the exams (F2), variety of the exams (F3), length of the exams (F4), complexity of the exams (F5), clarity of the exam instructions and question types (F6). Table 6 provides the mean scores and standard deviations of the students' viewpoints and beliefs about testing and evaluation. The applied questionnaire is a Likert-scale type. Considering the questions, number 5 represents 'strongly agree', and number 1 represents 'strongly disagree'. Number 3 indicates 'undecided'.

Table 6

Students' Perception of Testing Practices in terms of Compatibility, Frequency, Variety, Length, Complexity, Clarity.

	Themes	N	Mean	Std. Dev.
Low levels	Compatibility	48	2,9583	1,19544
	Frequency	48	1,8750	1,00398
	Variety	48	2,9219	1,23465
	Length	48	3,3438	1,23300
	Complexity	48	3,2031	1,020061
	Clarity	48	3,0104	1,14618
High levels	Compatibility	66	2,9848	1,00945
	Frequency	66	1,7982	0,81597
	Variety	66	2,9091	1,06576
	Length	66	2,5909	0,96540
	Complexity	66	3,5795	0,89782
	Clarity	66	3,6364	1,08679

The attitudes of the participants towards the assessment and evaluation system vary in different values. Considering the mean score of the data results, the lowest value is seen for the variety of the exam both in high and low levels. This value is 1.87 in low levels and 1.79 in high levels. On the other hand, this data changes in highest value. The highest mean for the low levels is 3, 34 which represent the length of the exam, however the highest mean for the high levels is 3.63 that represent the clarity of the exams. In this sense, low levels are mostly satisfied with the length of the exams, but high levels pleased with clarity of the exams.

Moreover, participants have a positive attitude about the complexity of exams and the usage of language (F6). The mean value of this section is 3.57 which have slight difference with the highest mean value. When it is considered the low level participants, these factors are a little lower value if it is compared to high level participants.

Standard deviation is a measure, which is used to evaluate the dispersion of a set of data values. If the data close to the mean, this refers to low standard deviation while a high standard deviation indicates that the data are separate out over a wider range of values. In the direction of this definition, based on the obtained data, standard deviation in high level is lower than low level participants. This indicates the consistency of the data.

4.3 Qualitative Results

Considering the fourth and the fifth research question, the data were gathered through an interview. The following part provides qualitative results regarding four different interview questions

4.3.1 The consequence of frequent testing. Through this interview question the participants express their viewpoints by discussing frequent testing by different types of exams. The responses were presented by the students for both levels. Students define their opinions about frequent testing, the effects of frequent testing on their progress and their language learning. As for the analysis of qualitative data table 7 illustrates the themes from the responses of interview questions about students' perception of frequent testing.

Table 7

The Result of Students' Perception about Frequent Testing.

Negative perceptions	Low	high
boring	9	8
mentally tiring	7	9
not interesting	7	1
leading to pressure	9	35
difficult to handle	3	3
unnecessary	4	6

Table 1 (cont. d)

leading to memorizing	4	7
time consuming	3	2
just involving grades	1	5
requiring heavy workload	5	14
difficult to handle	1	4
demotivating	2	4
reducing self confidence	-	3
Positive perceptions	Low	High
promoting my learning	1	5
boosting self confidence	1	4
motivating	-	1
balancing grades	-	2
helping me study regularly	-	1

The themes were divided into two distinct categories, which were positive and negative for low and high grades. According to the first category, it can be clearly seen that students' perception towards the exams, which is boring, is the most frequent theme that was gathered from both low and high level participant. Mentally tiring followed with 7 (low) and 9 (high) responses mentioned respectively. 35 of the high level participants declared that frequent testing leads pressure. Not interesting occurred 7 times in low level students' perception which is very low (1) in high levels. Quite few (3) statements by the students in both level suggest that they have difficulty to handle testing and assessment concept every week. Mostly high level participants' responses suggest that frequent testing leads to memorizing and unnecessary. Moreover, participant's comments also indicate that low levels need more time intervals than high level participants. 5 of the responses given by high level learners demonstrate that they just interested in their grades. Most of the high level participant mentioned that frequent testing requires heavy workload, 5 statements by the low level participants also share the same theme. On the other hand, quite a little low level participants feel too many exams are demotivating and reduce their self confidence while high level ones feel a lot in these statements.

Requiring heavy workload

[...] Being tested every week affects me in a negative way. In a week we have writing, listening, reading circle, and online speaking exam. This is really hard and boring. (low level)

[...] Apart from exams, we have assignments and presentations to do at the same time. It is too much for us. (high level)

Just involving grades

[...] I just memorize all the vocabularies for the exam, after that I forget everything. I am not doing it to learn something, just to get high grades. (low level)

[...] Every time I found myself studying my exams. It is a kind of race and we want to get good grades in this race (high level)

Boring, mentally tiring

[...] Instead of being tested every week there should be more fun activities or competitions not to get tired that much. (high level)

[...] It looks as if I come to school just for my exams. It is so boring and tiring. (low level)

[...] Just mid-terms and final exams are enough to assess, more than this is so tiring (high level)

Time consuming

[...] Being tested every week forces me to study and it takes a lot of time to practice and study everything. (high level)

[...] It is really time consuming, I don't have enough time to do something for myself. (low levels)

[...] It makes me socially passive because I cannot do anything except studying exams. (high levels)

Difficult to handle

[...] I sometimes do nothing except studying exams; it is so difficult to be tested every week. (high level)

[...] I also work out of school so it is difficult to handle for me to prepare exams. (low grades)

Reducing self-confidence

[...] Getting lower marks from the exams reduce my self confidence. (high level)

[...] Ever week we have exams and I don't think I study properly and I lose my self confidence. (low grade)

The positive perceptions of the participants about frequent testing mostly center around high level participants. Students' comments indicate that being tested every week promote their language learning. 4 statements by the high level students suggest that frequent testing boost their self confidence. Balancing grades appeared quite few (2) in high level students' positive perceptions. Motivating and help me study regularly were the lesser themes, yet still positive effect on frequents testing.

Promoting my learning

[...] I think being tested every week is beneficial and develop our language learning. (high level)

[...] Frequency of testing is important both for the institution to get regular feedback and learners to refresh our knowledge. (low level)

Boosting self confidence

[...] Frequent testing has a significant role for our language development and I fee self-confidence for this situation. (low level)

[...] I practice my lesson and by the help of exams and I feel self confidence. (high level)

Motivating

[...] Ever week we have different types of exams and I see lots of different questions this motivates me to study. (high level)

[...] I like learning English and being tested every week motivates me to study. (high level)

Balancing grades

[...] Thanks to different types of exams, I can increase my low grades. (high level)

[...] I always get low marks for writing assessment; thanks to frequent testing I balanced my grades with other assessment types. (high grades)

Helping me study regularly

[...] Since we have different types of exams every week, I study my lesson regularly. (high level)

[...] I don't like studying but exams provide me to study and practice my lesson regularly. (low level)

In accordance with the comments of participants through the second interview question, the analysis aims to investigate effects of progress about being tested every week

Table 8

The Effects of Frequent Testing on Students' Progress

	low	high
Negative effects		
loss of concentration	3	8
difficulty in learning process	20	34
anxiety	4	15
Positive effects	low	high
increased learning	13	34
higher self confidence	12	30
higher awareness	2	5

Table 8 shows the result of the consequence of frequency testing response of low and high level participants. In regards to the negative effects it can be clearly seen that difficulty in learning process is the most frequent theme that was gathered from the learners. 34 of the high level participants mentioned they have difficulty in learning process while this theme is 20 in low levels. 15 statements by the high level students suggest that they have exam anxiety which is little in low level. The least mentioned item for both levels of learners is loss of concentration which is 3 in low level and 8 in high level participants.

Anxiety

[...] Since this situation causes a lot of anxiety and we make more effort, it has negative effects on me. Rather than being tested every week, I prefer enjoyable assignments. (high level)

[...] I have a lot of thing to study and I feel anxious. (low level)

Loss of concentration

[...] We have difficulty in concentrating any of the topics that we should study. (low level)

[...] Every week we have the same routine, I can't concentrate on well to my studies. (high level)

Difficulty in learning process

[...] It takes me long hours to study English exams and it is difficult for me. (high level)

[...] It is difficult to study too many exams; it should be once and end of month. (low level)

On the other hand, participants have some positive comments on their progress about frequency testing. Most of the high level students stated, being tested every week increased and promote their learning. Considerable amount of the low levels supported the same theme. Not only high levels but also low levels expressed

that frequency testing and assessment provides higher self confidence. The positive perception of learners included higher awareness with 5 in high level while it is 2 in low levels.

Increased learning

[...] Since we are being tested every week I always revise my lesson and this promote my language learning. (low level)

[...] Improvements on my grades increase my learning level. (high level)

Higher self confidence

[...] This provides me self-confidence and development in language learning. (high level)

[...] Especially frequent speaking exams increase my self-confidence. (low level)

Higher awareness

[...] It is beneficial because I realize my missing subject and try hard to study more on those subjects. (high level)

[...] I see my development on my grades so it is beneficial for my progress. (high level)

Considering the third open ended question, the main purpose of the analysis was to evaluate the effects of being tested every week on language learning.

Table 9

Effects of Frequent Testing on Students' Language Learning

Negative Effects	Low	high
superficial learning	20	26
temporary learning	1	13
memorizing	3	4
Positive effects	Low	high
eagerness to learn	8	17
motivation to study	3	11

Table 9 shows the differences between the both levels of frequent testing towards their language learning. Superficial learning is the most frequent theme that was mentioned from both levels of students. Additionally, 13 statements by the high level participants suggested that frequent testing cause temporary learning while this theme stated only 1 in low level learners. Both levels of the participants believed that they memorize their topic or subjects rather than learning.

Superficial learning

[...] In fact, my aim is to pass the exams so I just study before the exams.
(high level)

[...] I memorize the structures that include exam topics to get higher marks.
(low level)

Temporary learning

[...] I study for the exams just because I have to, so it is not long lasting.
(high level)

[...] I could not memorize vocabulary after exams because I just study before the exams. (low level)

Memorizing

[...] Exams may be helpful for us to learn but it is not enjoyable and I just memorize grammar structures and some target words. (low level)

[...] I just focus on exams and I could not comprehend structures totally. I just memorize it. (high level)

Nearly half of the participants for both levels mention that being tested every week effect their language learning in positive way and support their language learning. 17 of the high level participants mentioned that they are eager to learn the language. Especially high level participants declared that being tested every week motivates them in their learning process.

Motivation to study

[...] There is an effect especially for my vocabulary knowledge and learning new vocabulary motivates me. (high level)

[...] I always learn new structures and words; it motivates me to study language. (low level)

Eagerness to learn

[...] The whole week we have different exam types so we have a chance to learn from our mistakes. (low level)

[...] I am always being tested for my language education. This is a kind of practice for me. (high level)

Motivation to study

[...] Apart from pop quizzes, other assessments have positive impact on my learning and I generally get higher marks this motivates me to study. (high level)

[...] Exams help me to learn more. (low grade)

4.3.2 Consistency of exams with teaching practices. For the last research question, the present study also investigates the students' perceptions of consistency of testing with teaching practices.

Table 10

Consistency of Testing with Teaching Practices

	Low levels	High levels
harder than the class activities	15	42
not compatible with the lessons	10	24
compatible with the lessons	13	20

Parallel to the last research question, the qualitative analysis of the interview which is examined the 124 students' perceptions about exams' content and their

compatibility with the lessons indicate that high level students with 42 responses declare exams are harder than the class activities. The same theme with 15 responses reported in low level participants. 10 statements by the low level students suggest that exams are not compatible with the lessons, whereas high level participants mentioned fewer. On the other hand, the positive perceptions of the students' comments towards consistency of testing with teaching practices included compatible with the lessons with 13 mentions in low levels and 20 mentions in high level learners.

Harder than the class activities

[...] Exams are compatible with the lessons but the content of the exam should be clearer and exam instructions should be easier. Also we cannot transfer our knowledge in such a short time. (low level)

[...] Use of language in the exam is harder than classroom language. It is sometimes difficult to comprehend the questions. (high level)

[...] Reading sections are too much and long in the exams. (high grade)

Not compatible with the lessons

[...] Actually, this is one of the crucial problems in our system. Our question types in classroom activities are different from the exam questions. (low level)

[...] Sometime I see unknown vocabulary in the exams. (high level)

[...] Especially listening part in the exams are not compatible with classroom practices. (high level)

Compatible with the lessons

[...] Exams are compatible with the lessons but they are harder than the class activities. Especially, reading and listening sections. (high level)

[...] When I see the similar types of questions in the exams that we practice in our classroom I feel secure. (low level)

[...] Our reading tasks in the exams are compatible with our course book unit.
(high grade)

In brief, the findings of the current study gathered from questionnaire and interview demonstrate that testing, assessment has a big impact on learners' accomplishments, and some visible aspects should be improved in accordance with the perceptions of students towards the content and compatibility of the assessment types and the frequency of testing.

Chapter 5:

Discussion and Conclusion

5.1 Discussion of findings for Research Questions

The purpose of this study was to investigate the effects of exams on students' language learning in the School of Foreign languages, in Izmir, Turkey. The study also aimed to investigate the learners' preferences about testing and assessment system and analyze the perception of students to comprehend how effective testing and assessment through their language development. In accordance with this purpose, data were collected both qualitative and quantitative methods; and mixed method research design was used to analyze the responses. The following sections discuss the findings of each result concerning the research questions.

5.1.1 Discussion about the general perceptions of test content by students and different level. Considering the first and second research question, the current study attempted to investigate the general perceptions of test content through the language learners and different levels. The result of the analysis demonstrate that the percentage of learners most favorite part is similar to the most helpful part of the exam which includes vocabulary and writing in highest percentages. The findings of the current study may be the proof of Brown's (2007) notion in such a way that, lexical items are basic to all of the four skills. Additionally, the highest percentage of vocabulary and writing in the analysis has the biggest effect on learners' final scores. Another finding in this section was the analysis of the highest mark that the learners get in the exams. As for the low levels the highest percentage demonstrates reading section while in high levels vocabulary part was the given response. As it is understood from the findings both receptive and productive skills were taught appropriately in the language school. One another findings aimed to investigate the most difficult part in the exams. Both levels, give responses as listening section, which may be insufficient class hours for the language learners or the content of the listening sections, may be difficult for them. On the other hand, in Preparatory School every other week learners are evaluated via extensive listening, which has an effect on their grades. This may cause anxiety for the learners. Similar to the findings

of Brindley and Slatyer (2002) who revealed some useful and constructive suggestions regarding confidential assessment task for the listening skills. The researchers claimed that testing listening skills is not easy to accomplish. The finding of difficulty in listening part affects the finding, which aims to analyze lowest grade that the students get from the exams. For the low levels, the results demonstrated that they have low grades for the listening section of the exams while in high levels grammar section was the given response. As for the given responses for high levels, grammar section may be found difficult because of the overloaded grammar syllabus. Finally, the most time consuming part in the exam was marked as writing section for the high levels while it was reading for the high levels. The exams for the high level students include two different tasks for reading section. This can also be an important reason why high level students have time management problem especially in reading section. Writing is the next most important part in this section, whereas vocabulary, listening and grammar sections can be said the least time consuming part and easy to accomplish.

To sum up, learners have no major difficulties in almost all skills, apart from listening. Moreover, it can be clearly expressed that learners found writing and vocabulary part more beneficial and constructive for their final grades and language development. On the other hand, both levels have difficulties and getting low grades for the listening part and they have time management problem especially with writing and reading sections.

5.1.2. Discussion about the findings of differences between the levels of students' general perceptions of test. The third research question attempt to assess the students' perceptions of testing practices in terms of compatibility with the teaching practices, frequency, variety, length and complexity of the exams, and clarity of the exam instructions and question types.

Considering the third and fourth research question, the lowest mean value as for the questionnaire of both levels demonstrates F2 item, which represent the frequency of the exams. As it is understood from questionnaire and interview results, both levels are not pleased with getting exams every week. Compatibility with the teaching practices is one of the problematic issues that should be considered especially low level participants. This may be related to washback effect. These

findings are parallel to the findings by Kirkpatrick and Gyem (2012) which argues to design any assessment system with the expectation that will help classroom teachers to keep a better balance between teaching and examination. Learners may have some difficulties when they face with unknown vocabulary or different types of questions in the exams. One can conclude from the findings that types of questions in the exam and classroom activities should be balanced. Not only receptive skills but also productive skills should be assessed with the similar types in classroom learning. According to the interview results, learners expressed that different kind of assessments do not improve their language development. The reason for that which is mostly for the high level participants is a kind of superficial learning which focuses on memorizing, lack of freedom in studying. On the other hand, learners all adopted to pass their exams. The results also demonstrate that a low level participant complains about the variety of the exams. Considering the interview findings learners claimed that they spent too much time to study their assessments. This can be tiring and affect their language acquisition in a negative way. It can be sometimes impossible to have a test that will accurately reflect every student's effort and knowledge. Some of the tests may not measure student learning as intended. Considering this issue, both levels perceptions of compatibility with the teaching practices was one of the problematic value. This situation may lead to stress and anxiety, which proves the data, gathered from interview and questionnaire. Rather than this, students need to come out of school to learn creative problem solving and study cooperatively. These findings are in accordance with the findings of Pekrun, Elliot, and Marier's study (2006) which analyze the emotions of language learners. Although the reading section was not marked as a favorite section in the first part of the questionnaire, length of each section in the exam and the length of questions and instructions rank in the highest mean value in the questionnaire. It can be clearly understood from those findings that most of the learners did not have time management problem during the exam. Moreover, difficulty level of the exams was fine for the learners. Some of the low level participants may have difficulty about the usage of the language or the instructions given in the exams. This can be related to the perceptions of low levels towards the language learning. Finally, for the low level participants' exam questions and the exam instructions were almost clear.

Considering the high level participants, when it is compared with the questionnaire and interview results, high levels had more dissatisfaction for the general perception of the test. As it was mentioned before, high level learners also complained about the frequency of the exams. They were not satisfied with being tested every week. According to their beliefs about the frequency of the exams, the learners just interested in their grades rather than learning. Being tested every week is not effective for them and it is just a temporary learning. This issue was also held in Zraggen's study (2009) which concluded that bi-weekly testing is more beneficial for learners than weekly testing. As for the high levels, the exam questions and the difficulty of the exams were not compatible with the classroom activities. Conversely, regarding the interview results, high level learners found exams harder than classroom activities and consequently they have time management problem in the exams. High level learners have almost the same value with the low levels regarding the variety of the exams. The collected data clearly show that students dislike the variety of the exams. On the contrary, this diversity decreases their confidence in using different skills. Unlike low levels, high levels are not satisfied with the length of the exams. This result can stem from the fact that reading sections in the exams for the high levels, there were two different tasks to accomplish. It can be suggested that the tasks of reading or the length of the questions should be prepared for the learners needs. As it is known, high level learners can comprehend the content of many items independently and they are good at analyzing. In this context, their general perception of the test in terms of difficulty of the exams and comprehending the questions are better than low levels. High level learners have positive attitudes towards the difficulty of the exams and clarity of the exam questions and instructions.

5.1.3 Discussion about the washback effects of classroom-based tests. The last research question of the current study aimed to explain the learners' perception of consistency of testing with teaching practices. There is diversity in learners' opinion on the questionnaire and interview question of their perception about the exam's content and their compatibility with the lessons. The collected data clearly show that both high level and low level learners mentioned exams are harder than classroom activities but high levels are persist with this comment. They believe that the use of language and the instructions in the exams are difficult. As it can be

inferred from the statements above, tests have some outcomes through its content and allocation of time. Especially high level learners think that they have time management problem during the exams.

On the other hand, learners have variable attitudes towards the compatibility of the exams. Both levels have some concerns about the compatibility of the exams with the lessons, while most of the low level participants have positive attitudes towards the compatibility between classroom exercises and tests. The reason is that some teachers' classroom practices are in relevance with the exam content. However, other teachers teach things that are different from the exam content. There were some studies to demonstrate some criterion whether the test influences is positive or negative. These facts are supported by one of the studies of Alderson and Wall (1993), which states if the teachers use tests for the students pay more attention to learning, it is positive influence of washback whereas if the teachers narrow the curriculum to make their students more successful with the fear of their students' poor result, it will be negative effects of washback. Moreover, Paris (1991) declared that standardized achievement tests encourage teachers and schools to spend valuable class time on what was measured at the expense of other school outcomes that were not tested. Another result that participants give response as a positive effect of consistency of testing with teaching practices is the effectiveness. Most of the language learners believe that the compatibility between the lessons and exams has positive effects on their language learning. Similarly Pan (1992) found that test motivate students to work harder to have a sense of accomplishment and thus enhance learning.

5.2 Pedagogical Implications

This study reveals some practical implications on the effects of testing on students' language learning in Preparatory School. The collected data revealed the strong and weak points of the assessment and evaluation system and the effects of classroom activities on testing. Considering both qualitative and quantitative data results, the students have negative attitudes towards the frequency of testing. Despite what is written in most of the literature in regards to benefits of the frequency of testing, the results of the present study indicated that students were significantly displeased with the number of assessment. As regards to these findings, it can be

recommended that assessments at the Preparatory School of Preparatory School should be minimized or allocate longer periods.

In addition, most of the participants in both levels have some concerns about the compatibility with the teaching practices. Most of them stated that exams are harder in comparison to the classroom activities and they declared this situation create stress and exam anxiety. In this sense, the testing committee with the teachers should reconsider those facts, which are mentioned below while preparing the exams.

Considering the points above, the findings of this study are significant for the effectiveness of testing for language learners. Implications and suggestions will hopefully provide benefit to not only the teachers at the Preparatory School, but also any other interested EFL instructors in the process of establishing a platform where students can be better evaluated.

5.3 Conclusion

The current study aimed to reveal learners' attitudes towards the assessment program at Preparatory School. The research also tried to indicate perceptions of learners' about the compatibility with teaching practices between testing and their attitudes about washback effect of classroom-based exams. The data collected through questionnaires and interview indicated that the frequency of the exams resulted in the exam anxiety and dissatisfaction for both high and low levels.

The responses of the learners as for the washback effect also include diversity. Both of the participant groups believed that exams are not compatible with the class-based teaching on behalf of the difficulty level, while low levels are more persistence about this statement.

To conclude, the study focus on a very specific and local context to find out the effectiveness of testing and assessment in terms of compatibility with the teaching practices, frequency, variety, length, and complexity of the exams and clarity of the exam instructions and question types. Therefore, it is significant to consider those results and provide possible solutions in order to reduce the problematic issues as much as possible.

5.4 Recommendation

The current study has some significant recommendations for the future studies. Based on findings, a further study could be conducted the students perceptions on the content of assessments can be compared by conducting document analysis of the assessments taken by the students. This would show the reliability of the exams and testing. It would be especially interesting to do research with each skill. In order to look for something in a particular way and interpret learner's behaviors, observations in the classroom atmosphere could be useful.

REFERENCES

- Ahmad, S., & Rao, C. (2012). *Research on Humanities and Social Sciences*. A Review of the Pedagogical Implications of Examination Washback, 2(7), 11-19.
- Aida, M., & Shaban, B. (2012). The Effect of Testing Frequency on Iranian Pre-intermediate EFL Learners' Language Achievement. *Journal of Academic and Applied Studies*, 2(10), 76-87.
- Allen, P., Fröhlich, M., & Spada, N. (1983). Department of Education. *The Communicative Orientation of Language Teaching. An Observation Scheme*, 23.
- Bachman, LF, & Palmer, A.S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford: Oxford University Press
- Bachman, LF, & Palmer, A.S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press
- Başol, G., & Johanson, G. (2009). Effectiveness of frequent testing over achievement: a meta analysis study *International Journal of Human Sciences* [online]. 6(2). Available from: <http://www.insanbilimleri.com/en>
- Bell, P. (2001). Content analysis of visual images. In *Handbook of visual analysis*, T. Van Leeuwen & C. Jewitt, (ed.) 10-34. Los Angeles, CA: Sage.
- Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1) 7-74.
- Brown, A. (1993). The role of test-taker feedback in the test development process: Test-takers' reactions to a tape-mediated test of proficiency in spoken Japanese. *Language Testing*, 10, 277-301.
- Brown, G.T., & Wang, Z. (2013). Illustrating assessment: how Hong Kong university students convince of the purposes of assessment. *Studies in Higher Education*, 38(7), 1037-1057.
- Brown, H.D. (2007). *Teaching by principles: an interactive approach to language pedagogy* (3rd ed.) New York: Pearson Longman
- Brown, H.D. & Abeywickrama, P. (2010). *Language assessment and classroom*

- practices* (2nd ed.) .White Plains, NY: Pearson Education.
- Cavanagh, R., Romanoski, J., Fisher, D., Waldrip, B. & Dorman, J. (2005).
Measuring student perceptions of classroom assessment. In Jeffery, R. and
Shilton, W. and Jeffery, P. (ed). *Creative Dissent: Constructive Solutions*, pp.
2-12. NSW: AARE Inc.
- Cheng, L. (1999). Changing assessment: Washback on teacher perceptions and
actions. *Teaching and teacher education*, 15(3), 253-271.
- Cheng, L. (2008). Washback, impact and consequence. In E Shohamy& N.H.
Hornberger (Eds.), *Language testing and assessment* (2nd ed., pp. 349-364).
New York: Springer
- Cumming, A. (2009).Language Assessment in Education: Test, Curricula, and
Teaching. *Annual Review of Applied Linguistics*,
doi:10.1017/S0267190509090084, 90-100.
- Davidson, C. (2007). Views from the chalkface: English language school-based
assessment in Hong Kong. *Language Assessment Quarterly* 4(1), 37-68.
- Davidson, C., & Leung, C. (2009). Current issues in English language teacher-based
assessment. *TESOL Quarterly* 43(3), 393-415
- Desheng, C., & Varghese, A. (2013). Testing and Evaluation of Language
Skills. *Journal of Research & Method in Education*, 1(2), 31-33.
- Dickins, P.R. & Germain,K. (1993) *Language Teaching: A scheme for Teacher
Education*.Oxford: Oxford University Press
- Duran, Ö. (2011). *Teachers' and Students' Perceptions about Classroom-Based
Speaking Tests and Their Washback* (Master's thesis) Available from Bilkent
University Library: Ankara
- Dustin, D. S. (1971).Some effects of exam frequency. *The Psychological Record*,
21, 409-414, from the PsycINFO database, No.1972-09643-001.
- Educational Testing Service (2009).*Guidelines for the Assessment of English
Language Learners*.1-27
- Freeman M,& Mathison. S. (2009).*Researching Children's' Experiences*. New
York: Guilford Press
- Harmer, J. (2012).*Essential teacher knowledge: Core concepts in English Language
Teaching*.England: Pearson Education

- Hossain, M., & Ahmed, K. (2015). Language Testing: an Overview and Language Testing in Educational Institutions of Bangladesh. *Advance in Language and Literary Studies*, 6(6), 81-84. (Doi:10.7575/aiac.all.v.6n.6p.80).
- Hughes, A. (2003). *Testing for Language Teachers*. (2nd ed.) Cambridge: Cambridge University Press.
- Johnson, K.E. & Vosmik, J. R. (2007). Assessing Student Learning: A Collection of Evaluation Tools. *Instructional Resource Award recipient*. Retrieved from <http://www.teachpsych.org/resources/Documents/otrp/resources/gottfried09.pdf>
- Keys, N. (1934). The influence on learning and retention of weekly as opposed to monthly tests. *Journal of Education Psychology*. 25(6).427-436.
- Kirkpatrick, R. & Gyem, K. (2012). Washback Effects of the New English Assessment System on Secondary Schools in Bhutan. *Language Testing in Asia*. 2(4)5-18 doi: 10.1186/2229-0443-2-4-5
- Kniveton, B.H. (1996). Student perceptions of assessment methods, *Assessment and Evaluation in Higher Education*, 21(3), 229-238
- Mussaway, S. A. J. (2009). *Assessment Practices: Student's and Teachers' Perception of Classroom Assessment* (Master's thesis). Available from CIE Master's Capstone Project
- National Research Council. (2000). *Tests and Teaching Quality: In-terim Report*. Committee on Assessment and Teaching Quality, Board on Testing and Assessment. Washington, DC: National Academy Press
- Pawlak, M. & Klimczak, E. W. (2015). *Issues in Teaching Learning and Testing Speaking in a Second Language* New York: Springer
- Pekkanlı, İ. (2010). Educational Policies: Traditional Foreign Language Classroom Assessment in Turkish High Schools. *Turkish Studies*, 11(3), 465-476. doi: 10.1080/146833849.2010.505725
- Qiuxian, C., Margaret, K., Val, K., & Lyn, M. (2013). Interpretations of formative assessment in the teaching of English at two Chinese universities: a socio cultural perspective. *Assessment & Evaluation in Higher Education*, 38, 831-846.

- Sato, T., & Ikeda, N. (2015). Test-taker perception of what test items measure: a potential impact of face validity on student learning. *Language Testing in Asia*, 5(10), 2-16.
- Şenel, E. & Tütüniş, B. (n.d.). *The Washback Effect of Testing on Students' Learning in EFL Writing Classes*.
- Shim, K. N. (2009). *An investigation into teachers' perceptions of classroom-based assessment of English as a foreign language in Korean primary education* (Doctoral dissertation). Retrieved from <https://ore.exeter.ac.uk>
- Struyven, K., Dochy, F. & Janssens, S. (2004). Students' perceptions about evaluation and assessment in higher education. *Assessment & Evaluation in Higher Education*, 30(4), 331-347.
- Struyven, K., F. Dochy, & S. Janssens. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment & Evaluation in Higher Education* 30(4) 325–41
- Tsagari, D. (2007). *Review of Washback in Language Testing: How has been done? What more needs doing?* Lancaster University, UK
- Xiao, Y. & Carless, D.R. (2013). Illustrating Students' Perception of English Language Assessment: Voice from China, *RELC Journal*. 44(3) 319-340. doi: 10.1177/0033688213500595
- Zraggen, F. D. (2009). *The Effects of Frequent Testing in the Mathematics' Classroom* (Unpublished doctoral dissertation). The Graduate School University of Wisconsin-Stout Menomonie, WI.

APPENDICES

A: QUESTIONNAIRE SAMPLE

STUDENT QUESTIONNAIRE

Dear Students,

I am studying on my Master's degree in English Language Teaching Program at Graduate School of Educational Science, in Turkey. The aim of my thesis is to investigate the effects of exams on students' language learning in school of foreign language in Izmir, Turkey.

This questionnaire has been prepared to examine what extents do each of the exams help learners to comprehend language in Preparatory School of Gediz, in Izmir. The questionnaire consists of four sections. The first part aims to obtain information about personal profile, while the other parts attempt to get general perceptions of test content and your perception about the testing practices in terms of frequency of the exams, compatibility with the teaching practices, clarity, length , complexity, and variety of the exams.

Your answers might affect the results of an academic study. In line with the findings, suggestions will be made and current practices will be re-evaluated for improvement. Your answers will be used only for this research and will be kept confidential until they are disposed of.

Thank you for your contribution and collaboration.

Gamze Taşlı

The Department of English Language Teaching

MA student at Bahçeşehir University

gamzetasli@gmail.com

SECTION 1: STUDENT QUESTIONNAIRE

I want to participate: _____ (You may use my responses for research only)

I do not want to participate: _____

Age:

Gender: Male () Female ()

Level of the student: A1 A2 B1 B1+ B2

Please answer the questions

	Vocab.	Reading	Listening	Writing	Grammar
Q1. What is your favorite part in the exams?					
Q2. Which part of the exam do you think is the most helpful and has the biggest effect on your final result?					
Q3. From which part do you generally take the highest point in exams?					
Q4. Which part of the exam is the most difficult?					
Q5. From which part do you take the lowest point in the exams?					
Q6. Which part of the exam takes the most of your time in exams?					

SECTION 2

Please read the statements carefully and indicate the degree to which you agree with the statements listed below and tick (✓) the appropriate responses.

5 (strongly agree) 4 (agree) 3 (neutral) 2 (disagree) 1 (strongly disagree)

ITEMS	1	2	3	4	5
1. I like being tested every week.					
2. Being tested every week motivates me.					
3. Being tested every week enhances my language learning ability.					
4. Being tested every week enhances my self-confidence.					
5. Questions in the exams are compatible with teachings.					
6. The difficulty level of exam questions is the same with the classroom activities and worksheets.					
7. Exam questions are compatible with exercises and activities in courses.					
8. The difficulty of the exam questions is compatible with the classroom activities.					
9. Exam questions are understandable.					
10. The exam instructions in the exam are clear.					
11. Question texts and content of the exam is appropriate for my level.					
12. The types of the questions in the exam are compatible with the classroom activities.					
13. The length of the exam is compatible with the duration.					
14. Any part of the exam length is compatible with the duration.					
15. The length of the exam questions is compatible with the duration.					
16. The length of the instructions given in the exams is appropriate.					
17. The content of the questions in the exams are understandable.					
18. The level of English used in the exams is appropriate.					
19. The use of language for the instructions is understandable.					
20. The level of the use of English and distracters is appropriate.					
21. I like the variety of exams (exit, achievement, writing portfolio, online speaking, listening assessment)					
22. The variety of exams (exit, achievement, writing portfolio, online speaking, listening assessment) motivates me.					
23. The variety of exams (exit, achievement, writing portfolio, online speaking, listening assessment) enhance my learning.					
24. The variety of exams (exit, achievement, writing portfolio, online speaking, listening assessment) increase my confidence in using different skills.					

B: INTERVIEW QUESTIONS

Please respond as honestly and clearly as possible.

1. What do you think about being tested every week?
2. How does being tested every week affect your progress?
3. How does being tested every week affect your language learning?
4. What do you think about the exams' content and their compatibility with the lessons?

C: CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Taşlı, Gamze

Nationality: Turkish (T.C.)

Date and Place of Birth: August13, 1982, Söke ,TURKEY

Work Phone: 0090 232 35500

Mobile Phone: 0 533 212 9889

E-mail: gamzetasli@gmail.com, gamze.tasli@gediz.edu.tr

EDUCATION

Degree Graduation	Institution	Year of
MA	Bahçeşehir University	2016
BA	Uludağ University	2006

WORK EXPERIENCE

Year	Place	Enrollment
2006-2009	Bursa Kültür Collage English	Instructors
2009- Present	Gediz University	Instructors

FOREIGN LANGUAGES

Advanced English

CERTIFICATES

CLIL for Universities, Celt Academy, 2013

PUBLICATIONS

1. Taşlı, G. & Singör, T. (2011) “ The Impact of Sentence Production on Achievement Outcomes” . *Teacher - Research Studies at Foreign Language School: Inquiries from Teacher Perspectives*. (1)131-145

2. Taşlı, G.(2012) “A Successful Student Profile: From Student’ Perspective” .
Teacher-Research Studies: Inquiries from Teachers’ Perspective, (1), 169-175

HOBBIES

Travelling, Reading, Trekking