HOW WELL DO VARIOUS ASSESSMENT PRACTICES OF THE
PREPARATORY YEAR ENGLISH PROGRAM PREDICT THE SUCCESS OF
STUDENTS FOR TOEFL ITP (INSTITUTIONAL TESTING PROGRAM)?

A CASE STUDY AT TOBB UNIVERSITY OF ECONOMICS AND
TECHONOLOGY DEPARTMENT OF FOREIGN LANGUAGES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ŞEYMA DOĞRU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF ARTS
IN
THE DEPARTMENT OF FOREIGN LANGUAGE EDUCATION

SEPTEMBER 2013

Approval of the Graduate School of Social Sciences

_____

Prof. Dr. Meliha ALTUNIŞIK

Director

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

_____

Assoc. Prof. Dr. Nurten BİRLİK

Head of Department

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 Arts.

_____

Dr. Deniz ŞALLI ÇOPUR

Supervisor

Examining Committee Members

Y. Doç. Dr. Bilal Kırkıcı (METU, FLE)   _____

Dr. Deniz ŞALLI ÇOPUR (METU, FLE)   _____

Dr. Hale Kızılcık (METU, DML)   _____

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: Şeyma DOĞRU

Signature:

ABSTRACT

HOW WELL DO VARIOUS ASSESSMENT PRACTICES OF THE
PREPARATORY YEAR ENGLISH PROGRAM PREDICT THE SUCCESS OF
STUDENTS FOR TOEFL ITP (INSTITUTIONAL TESTING PROGRAM)?
A CASE STUDY AT TOBB UNIVERSITY OF ECONOMICS AND
TECHONOLOGY DEPARTMENT OF FOREIGN LANGUAGES

DOĞRU, Şeyma

M.A., Department of Foreign Language Education

Supervisor: Dr. Deniz ŞALLI ÇOPUR

September 2013, 106 pages

Students studying in the preparatory program at TOBB ETU are assessed with various assessment practices (placement test, midterms, quizzes, and readers) and required to pass TOEFL ITP test to start their undergraduate education in their departments. However, how well these assessment practices predict students' performance in TOEFL ITP is a question that needs to be investigated. Therefore, the purpose of this study is to find out the relationship among in-house assessment practices and TOEFL ITP. Moreover, it aims to investigate both the instructors' and students' opinion on the effectiveness of these assessment practices on TOEFL ITP. Forty-eight instructors and 421 students filled in a half open-ended questionnaire,

and the scores of 337 students were analyzed using Pearson Product Moment Correlation and Multiple Regression. The results showed that a higher correlation exists between midterms and TOEFL ITP, a situation which both instructors and students agreed on. In addition, low but still significant relationship was found between all the quizzes except one and TOEFL ITP. For the placement test, while no significant correlation was found for beginner and elementary levels, a low but significant correlation exist in intermediate and upper-intermediate levels. Lastly, even though a positive correlation is found between the average of reader exams and TOEFL ITP, more than half of the participants still think that they are the least efficient assessment practice.

,

ÖZ

TOBB EKONOMİ VE TEKNOLOJİ ÜNİVERSİTESİ YABANCI DİLLER

BÖLÜMÜNDEKİ FARKLI ÖLÇME ARAÇLARININ TOEFL ITP'yi YORDAMA

GEÇERLİLİĞİ

DOĞRU, Şeyma

Yüksek Lisans, Yabancı Dil Eğitimi Bölümü

Tez Yöneticisi: Dr. Deniz ŞALLI ÇOPUR

Eylül 2013, 106 sayfa

TOBB ETÜ Yabancı Diller Bölümünde İngilizce Hazırlık programına devam eden öğrenciler farklı ölçme yöntemleriyle değerlendirilir. Bu ölçme araçları düzey belirleme sınavını, ara sınavları, kısa sınavları ve okuma sınavlarını içerir. Yine bu öğrenciler hazırlık eğitimi sonunda bölümlerine başlayabilmeleri için TOEFL ITP sınavından gerekli olan puanı almalılardır. Ancak bu ölçme metotlarının dönem sonunda yapılan TOEFL ITP sınavını yordama geçerliliği sorgulanmalıdır. Bu sebeple, bu çalışma TOBB ETÜ'de hazırlanan ve uygulanan değişik ölçme araçlarının TOEFL ITP ile olan ilişkisini öğrenmeyi amaçlamaktadır. Buna ek olarak, yine bu çalışma iki farklı açık uçlu anket ile bu ölçme araçlarının etkisi hakkında hem öğrencilerin hem de okutmanların fikirlerini almayı hedefler. Kırk sekiz okutman ve 421 öğrenci ankete katılmış, 337 öğrencinin sınav sonuçları Pearson Product Moment korelasyon ve çoklu regresyon ile analiz edilmiştir.

Çalışmanın sonunda hem öğrencilerin hem de okutmanların umduğu gibi ara sınavlar ile TOEFL ITP arasında yüksek düzeyde pozitif yönlü bir ilişki bulunmuştur. Buna ek olarak tek bir kısa sınav haricindeki tüm kısa sınavlar ile TOEFL ITP arasında düşük ancak yine pozitif yönde anlamlı bir ilişki bulunmuştur. Düzey belirleme sınavı ile TOEFL ITP arasında başlangıç ve orta seviyelerde herhangi bir ilişki bulunmazken yüksek seviyedeki gruplarda pozitif yönde anlamlı bir ilişki bulunmuştur. Son olarak, her ne kadar öğrencilerin okuma sınavlarından aldıkları notların ortalaması ve TOEFL ITP arasında pozitif yönde anlamlı bir ilişki bulunsa da katılımcıların yarısından çoğu bu sınavların TOEFL ITP başarısına olan etkisinin en az olduğu görüşündedir.

Anahtar Kelimeler: Yordama Geçerliliği, Başarı Testleri, Korelasyon,

To My Family...

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER I

INTRODUCTION

This present study is designed to investigate the predictive power of the assessment practices in the Department of Foreign Languages at TOBB Economics and Technology University. Moreover, it aims to investigate both the students' and instructors' opinions about the predictive power of these assessment practices. In this chapter, the rationale and background of the study, its significance and purpose, background information about the Department of Foreign Languages, the tests carried out in this department, research questions and key terminology are presented.

1.1 Rationale, Background and Setting

Testing, the evaluation of an individual's proficiency in order to learn what level of knowledge or skill has been acquired (Lado, 1961), has been an issue of attraction for years. Moreover, in many institutions, language tests serve not only as an assessment tool, but also a teaching tool by the positive backwash effect they have on the curriculum. Bachman (1990) explains the fundamental purpose of them as collecting information for making decisions not only about test takers but also about the program. Likewise, Aiken (1997) explains the usage of various tests as to choose applicants for jobs or different contexts, to place people in educational settings, to guide, promote and rotate students for educational programs. Therefore, testing and evaluation is a very important research area not only for the testers or the test takers but also for researchers who have an interest in curriculum and instruction.

When language tests are taken into account, their aims seem to overlap as collecting information so that decision making process can take place. However, it is a very well-known fact that there are obviously several differences among them. As put forward by Bachman (1990), they differ from one another in terms of their purpose, frame of reference, design, scoring procedure and method. Alderson, Clapham and Wall (1995) propose broader categories for language tests as placement, progress, achievement, diagnostic and proficiency. When these different types of exams are examined, the first one, also to be used in this study, is achievement tests, an indispensible part of any institution that aims to assess the students' learning. They can be defined as the tests aimed to gather information during or at the end of a course in order to examine whether any progress takes place or not in terms of teaching objectives (McNamara, 2000), or as the tests carried out to learn how well students have achieved the instructional goals of a course or a program (Brown, 1996). By the help of such tests, curriculum can be renewed, revised or developed according to what needs to be improved and what has been achieved.

The second type of exam to be used in this research is the placement exam, carried out on the first day of the term. It aims to assess whether a student has the necessary language abilities for his/her studies in their departments, in other words, whether they are linguistically competent or not. While some institutions prefer to use international exams, the others including TOBB ETU, use in-house ones. As placement exams are conducted on the very first day of the term, they are not based on a specific language teaching curriculum but rather the main focus is on how much is known and if it is enough for a specific purpose.

The last type of the exam carried out at TOBB ETU is TOEFL ITP (Test of English as a Foreign Language), a paper-based test, which requires the students to answer 140 multiple choice questions on listening, structure and reading. These tests offer colleges and universities, English language learning programs and other

organizations the opportunity to administer a convenient, affordable and reliable assessment of English language skills.

Different kinds of tests including placement tests, admission tests, language aptitude tests and proficiency tests can all be explored for their predictive validity (Brown, 2004; Davies, 1990). By predictive validity, the possible future success or failure of a test taker can be foreseen (Bachman, 1990; Hughes, 2003) and up to now, there have been many researchers who have carried out studies in order to explore the predictive validity of tests. However, this study will differ from the previous ones reviewed as there is more than one variable investigated to predict a final assessment test.

1.2 Significance and Purpose of the Study

The related literature shed light on the problem that most studies are conducted just to investigate the relationship between two variables. However, the number of studies examining more than two variables is limited. Moreover, there is not any study carried out at TOBB ETU in order to examine the relationship between many achievement practices. In addition, what both the instructors and students at this institution think about those assessment practices in terms of their predictive power has never been investigated. Therefore, this present study addresses abovementioned problems as well as the untouched issues. Moreover, it is a great interest to the administration, to the members of testing and curriculum unit, all the instructors, students at the Department of Foreign Languages at TOBB ETU and lastly anyone who is interested in examining the power of assessment practices. More specifically, the aims of this study are:

1. To find out the relationship between the various types of achievement tests and TOEFL ITP at different levels of the Department of Foreign Languages, TOBB ETU.

2. To find out both the students' and instructors' perception towards the assessment practices and their relationship with TOEFL ITP

The context of study, TOBB ETU, is a private university, which was founded in 2003. Since then, there has been no study carried out in order to evaluate the predictive validity of the assessment practices. Therefore, a study to investigate the relationship between the variables and the predictive power of the tests used in the Department of Foreign Languages is extremely needed. The results of the study shed light on the effectiveness of the tests as the higher predictive validity of an assessment practice, the better inference can be made related to students' achievement on the following tests. In other words, if the predictive validity of the in-house achievement tests is found high, it might be an indication of the success of testing unit, more specifically the tests designed by its members. However, a change in the achievement practices might be necessary if low predictive validity is to be found. Therefore, all of these possibilities are dealt with the help of this present study. Moreover, both the instructors and the students of TOBB ETU present valuable information related to the assessment practices used and their relationship with TOEFL ITP.

1.3 Background Information about Department of Foreign Languages, TOBB ETU

At the Department of Foreign Languages in TOBB ETU, there are two periods in an academic year, but the second period is divided into two semesters. The first period consists of 15 weeks. While the first semester of the second period is 15 weeks, the second one lasts nine weeks. There are in total of 12 quizzes, ten midterms, and eight reader exams (which require students to read a predetermined book and answer related questions). All of these achievement practices are of great importance, since

4

only the students whose average score is 65 or above can take the TOEFL ITP to continue their education in their departments.

In Figure 1 below, a detailed explanation of the steps that students go through in the Department of Foreign Languages at TOBB ETU is provided.

```
                    ┌────────────────────────────┐
                    │ Department of Foreign Language │
                    │          TOBB ETU             │
                    │      PLACEMENT EXAM           │
                    └────────────────────────────┘

   ┌──────────────┐   ┌──────────────┐   ┌──────────────┐
   │ 49 & BELOW   │   │ BETWEEN 50-64│   │ 65 AND ABOVE │
   │ AF (BEGINNER)│   │ A (ELEMENTARY)│  │              │
   └──────────────┘   └──────────────┘   └──────────────┘

        ┌──────────────┐  ┌──────────────┐  ┌──────────────┐
        │ TOEFL ITP    │  │ TOEFL ITP    │  │ TOEFL ITP    │
        │ 450 & BELOW  │  │ 499-450      │  │ 500 & ABOVE  │
        └──────────────┘  └──────────────┘  └──────────────┘

            ┌──────────────┐  ┌──────────────┐  ┌──────────────┐
            │ B            │  │ C            │  │ PASS         │
            │ (PRE-        │  │ (INTERMEDIATE)│ │ TO THEIR     │
            │ INTERMEDIATE)│  │              │  │ DEPARTMENT   │
            └──────────────┘  └──────────────┘  └──────────────┘
```
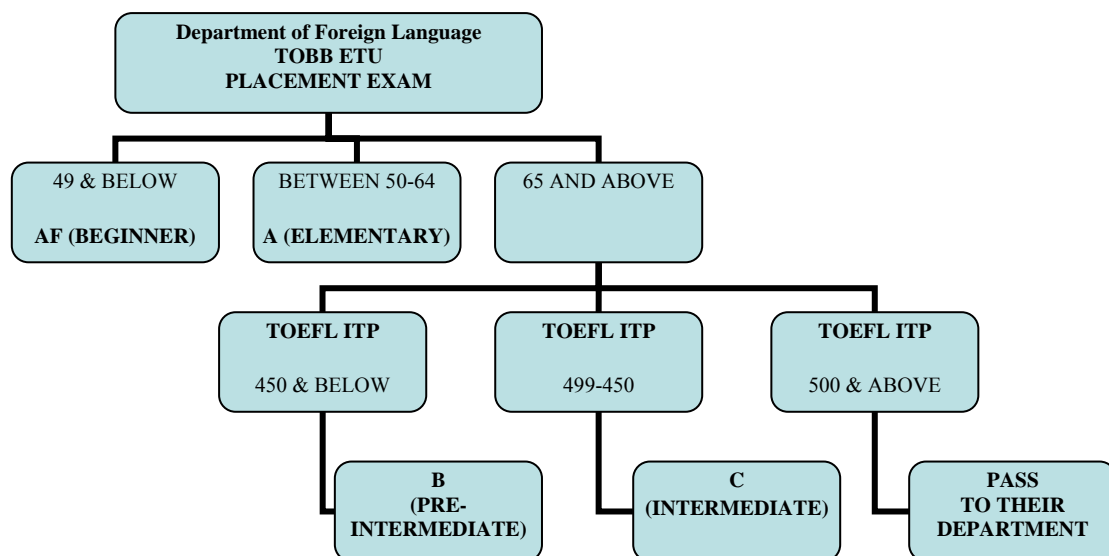
Figure 1: The Scheme of the TOBB ETU Preparatory Classes

In this study, for C level (upper-intermediate students), three midterms, three quizzes and three readers in the first term are taken into account as they are only expected to have their education till December; however, for the rest of the students, all the achievement tests were analyzed.

Right after the placement test results, two of the levels, beginner level (AF) and elementary level (A), are composed by choosing the students whose placement exam scores are below 65. The other students with a score above 65 are given a chance to take the TOEFL test. If they score 500 or more in this exam, they are not required to study in the preparation class. The ones getting below 500 are divided into two groups; while the ones having a score between 450-499 are put into upper

intermediate level (C), the others getting below 449 are put into pre-intermediate level (B).

In this study, both the scores of C Level students who took TOEFL ITP at the end of December and the scores of the students who took TOEFL- ITP at the beginning of July from the other three levels- AF, A and B Levels, were examined. More specifically, the relationship between the placement exam they took on the very first day, quizzes, midterms, reader exams and TOEFL-ITP was examined for all these four levels.

In Figure 2, both the number of each testing tool used to assess the performance of students in levels AF, A and B in the first, second and third term, respectively and the percentage of them are described in a detailed way.
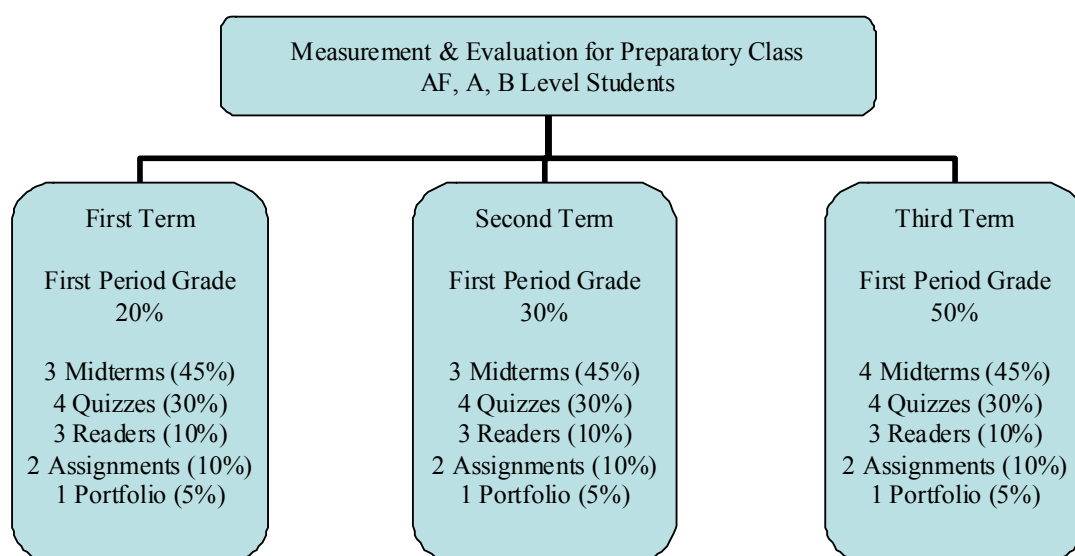
Figure 2: Measurement & Evaluation for Preparatory Class Students of A, AF and B Levels in the First, Second and Third Term

However, the percentages are different for C Level students as they finish the preparatory program in December. Therefore, in Figure 3 below, both the number of each testing tool and their percentage are described in a detailed way.
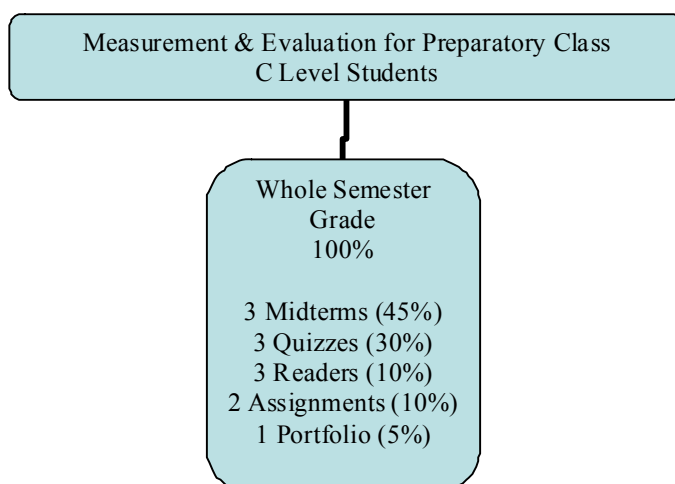
Figure 3: Measurement & Evaluation for C Level Preparatory Class Students

1.4 Assessment Practices at the Department of Foreign Languages, TOBB ETU

1.4.1 Assessment Practices Prepared, and Conducted by the Testing and Evaluation Unit

There are five different types of tests used in the Department of Foreign Languages at TOBB ETU. All of the tests, except TOEFL ITP, are prepared by the instructors working at the testing and evaluation unit. To begin with, placement exam (60 structure, 40 reading and 25 listening questions, total of 125 questions) was prepared by five instructors, one of whom is a native speaker of English. First, the questions were prepared by taking the distribution of the topics along with different test specifications into account, and then, all the testers gave feedback to each other.

The midterms, quizzes and readers are prepared by a single person responsible for each level. However, everybody in the testing unit is required to read the assessment practices of other levels and give feedback. The instructors working at the curriculum department also go over each test to be sure that everything included in the assessment tests has also been covered in class. The midterms consist of 20 questions in three different skills; listening, reading and writing. Besides, questions related to grammar points and vocabulary items are asked. On the other hand, the quizzes have four parts with 10 questions each, but the listening and writing questions are asked interchangeably. That means when a quiz has a writing part, the next quiz will have a listening part. In addition, there is one speaking quiz in each semester, which requires students to talk about themselves, a picture that they have chosen and a topic that is randomly chosen. Lastly, the students are expected to read a pre-determined short story or a book, and then they take an exam including comprehension and vocabulary questions related to it. This last type of exam is called reader exam and the students take eight of those in total, three of which are conducted in the first semester.

1.4.2 Assessment Practices Prepared, and Conducted by the Curriculum and Development Unit

There are two different types of assessment practices prepared by the curriculum and development unit in the Department of Foreign Languages at TOBB ETU. The first one is seven portfolio tasks which include writing, extensive reading, integrated research assignments, story writing and creating videos about different topics. The second one is assignment and project grades. The instructors are expected to give this grade based on the percentage of weekly homework carried out by students. These pieces of homework consist of vocabulary, reading, and workbook activities.

1.5 Research Questions

The study focuses mainly on the correlation between the assessment practices prepared, and conducted by the Testing and Evaluation Unit of TOBB ETU and TOEFL ITP. As illustrated in Figure 2 and 3, the assessment practices include readers, midterms, quizzes and placement exam results. Therefore, for this study, the relationship between all these assessment practices and TOEFL ITP along with their predictive power is investigated. In addition, both the students' and instructors' opinions on the topic were taken. The research questions of this study are as follows:

1. How well do the different types of achievement tests including midterms, quizzes, readers and placement exam grades predict the TOEFL ITP scores of students studying at the preparatory program at TOBB University?
2. Is there a significantly positive relationship among the different types of achievement tests including midterms, quizzes, readers and placement exam grades and TOEFL ITP scores of students studying at the preparatory program at TOBB University?
3. What are the students and instructors perceptions towards the relationship between all these assessment practices and TOEFL ITP?

1.6 Definition of Terms

Predictive validity: An objective measure referring to the degree to which a test can predict the possible future success or failure of a test taker (Bachman, 1990; Hughes, 2003). In this study, predictive validity refers to the relationship between several achievement tests and TOEFL ITP in terms of their predictive power.

Placement Tests: Tests used to determine the selection of students that will study in an English preparation school and to judge whether candidates are ready to implement a task or an activity (Enginarlar, 2002) and in this study, they refer to the

test carried out on the first day of English preparatory year to place students according to their level. There are three different parts in this test, namely listening, structure and reading and the number of question in those parts are 25, 60 and 40, respectively.

Achievement Tests*:* Tests aimed to measure the amount of learning in a prescribed domain by being in line with explicitly stated objectives of a program (Henning, 1997). In this study, achievement tests refer to 10 midterms, 12 quizzes and eight readers.

Midterm: Midterms are also examples of achievement tests and in this study; they refer to ten achievement tests. Seven of them include some of the areas and skills, namely grammar, vocabulary, structure, reading and writing, and in each midterm there are twenty questions. However, the format of the last three midterms is different. Like TOEFL ITP, in the last three midterms, there are also 50 listening, 40 structure and 50 reading questions.

Quiz: In this study, quizzes refer to types of measurement in order to check student acquisition of skills or knowledge following appropriate exposure to the material. There are 12 quizzes carried out in Foreign Languages Department. While three of them assess students' speaking skills, listening, reading, structure and vocabulary questions are asked along with a writing task in the other nine.

Reader or Reader Exam: In this study, readers refer to eight short quizzes which include both comprehension and vocabulary questions related to the short books chosen. The students are expected to read the assigned books before taking the reader quiz and answer the question about them. As it is a closed-book exam, the students are expected to learn the content**.**

CHAPTER II

LITERATURE REVIEW

In this chapter, related studies in the literature are presented. The most relevant ones to the present study carried out both in Turkey and all around the world are given under specific titles in connection with the variables. Not only the ones including the data coming from university level students but any study dealing with different levels from primary to high school are exemplified. All the studies both related to the local and global context are presented chronologically.

2.1. Predictive Validity

There are many different qualities of tests which need to be taken into account while constructing, administering and more importantly interpreting them. The first one is the validity issue. For any tests, the scores should not be affected by any factors other than the ability that is being tested (Bachman, 1990). The tests should only measure the areas that are being covered. There are different types of validity such as face validity, construct validity, concurrent validity and predictive validity which is the focus of interest in this present study,.

Predictive validity has always been a topic of interest for many researchers. As it is regarded as a very strong measure of statistical validity, it is used in many institutions and by many researchers. In predictive validity, while high correlation shows that the selected procedure works in the right way, a low correlation indicates that the approach needs to be changed or adapted somehow (Shuttleworth, 2009).

11

2.2 The Relationship between Various Assessment Tools and Academic Achievement and their Predictive Power

2.2.1 Studies at Global Context

An example of a predictive validity study was carried out by Shulruf (2011). He investigated the assessment tools used to select medical students for a six-year medical program. As three different admission tools were used in Australia, to what extend the scores on each of these tools predict the academic achievement was the main research question to be answered. The admission tools were admission GPA (Grade Point Average), meaning the prior academic achievement measured by grade point average, the average UMAT score (Undergraduate Medicine and Health Sciences Admissions Test) and lastly a structured interview with two interviewers lasting 25 minutes. Like previous studies, gender, age or ethnicity was not also taken into account. Data were collected from 324 students during the consecutive years of 2005, 2006, 2007 and analyzed by using regression analysis. Among the participants, 157 were male ($p = 49\%$) and 167 ($p = 51\%$) were female. At the end of the study, admission GPA was found to be best predictor of academic achievement in each year. However, the prediction of the other two admission tools was low in comparison to admission GPA. Lastly, the scores on each admission tools were also compared and relatively low or no correlation among them was found. This suggest the little resemblance and overlap in what they are intend to measure and necessity for all of them to be used in the admission process.

2.2.2 Studies at Local Context

A different study was carried out in Turkey in 2011 by Zırhlıoğlu and Atlı. In their study, the scores of student selection exam for the university education (OSS) and the ability test required for Physical Education and Sport departments and GPAs (Grade Point Average) were used. More specifically, by using a stepwise regression method, these two assessment tools and students' high school GPAs were examined

12

in relation to their academic success in university. The data were collected between the years of 2005 till 2008 in Yüzüncü Yıl University, Van, Turkey. At the end of the study, a low correlation was found between academic success and student selection exam, GPA and the ability tests. That might be an indication of various variables affecting students' academic success. Therefore, it might be better to examine external factors which may interfere to the results of the study.

2.3 The Relationship between Nationwide Exams and Academic Achievement and Their Predictive Power

2.3.1 Studies at Global Context

A study which analyzed the predictive validity of SAT for high school grades from 1976 to 1985 was carried out in 1989 by Morgan with a longitudinal design. Therefore, it can also show the degree of the predictive power of SAT scores during those nine years of period. Even if at the end of the study, it was found out that there was a decline in the predictive validity of SAT scores, the degree to which this decline resulted from the changes in the SAT during those years is still a topic to be investigated.

Bronner, Kennet-Cohen and Oren (1999) investigated the predictive validity of the process through which the candidates attend higher education in Israel with respect to either their cumulative grade point average (CGPA) at high school or their freshman grade point average (FGPA). They used multiple regression analysis to examine the predictive power of these two average points. The students from six Israeli universities in the school years of 1991-92 and 1992-93 participated in the study; while 24.969 of them had FGPA, the rest 16.731 reported their CGPA in total of 41.700. The results of this study showed a higher predictive value in the samples of students with CGPA than those with FGPA. The explanation behind such a difference was explained as the difference between these two criteria.

The validity of ACT (American College Testing) test scores and self-reported high school grades for predicting the grades of specific college freshman courses was examined by Noble and Sawyer (2004) for four years. For this purpose, students records containing ACT scores, high school average GPA and self-reported high school grades in different areas such as English, Math and social studies were collected. For each area, sub-categories were formed and multiple regressions were used for predicting each course and its sub categories. For some courses such as English, multiple regression analysis showed similar results across its sub-categories, but for some other courses such as Math, it was not the case. Accounting regarded a sub-category of Math showed the highest prediction of ACT tests scores. They further suggested the use of these data for benefits of students. If a student with low predicted chances of success in any of the courses or its sub-categories, he might be advised to pursue a different path; likewise, a student with high predicted changes of success can be guided to take up those courses.

There are numerous studies carried out in order to find out the relationship between two different variables, one of which was carried out in 2005 by Kobrin et al. The main purpose of the study was to find an answer to the question of how well the SAT test can predict first-year college performance. In order to get the files containing students' first year performance data such as their course work, grades and GPA, 110 institutions were contacted. These data were matched to the SAT scores in College Board databases. The study compared the correlation of predictors, in this case SAT scores and GPA. After examining the data of 1.316 participants, it was found out that there is a positive correlation between these two, meaning the students who performed better on the SAT had a higher GPA. However, as it is mentioned before, this study also investigates the correlation between two variables. On the other hand, in this present study, more than two variables are taken into account which creates a better picture of the area being investigated.

Another study again taking only two variables into account was carried out in 2005 by Donald and Christina in order to examine the validity of a different type of exam called Pharmacy College Admission Test (PCAT) for predicting grade point averages of students in pharmacy programs. The data were collected from 11 colleges and schools of pharmacy, and correlation analysis used to determine the validity of the test for predicting GPAs, showed that PCAT test scores and GPAs were positively correlated.

Even though there are many studies showing that SATs are good predictors of students' academic performance or success measured by GPA, there are not many of them investigating the new SAT introduced in 2006. Therefore, the study carried out by Carty and Armstrong (2008) found an answer to the questions whether the new SAT test had predictive validity in predicting undergraduate success and whether there was a difference between the new and the old one in terms of predicting undergraduate students' academic performance. Moreover, the predictive power of high school grades to the undergraduate success was also the focus of the study. The data were collected from 3.981 and 4.589 students enrolled in a research institution in years of 2001 and 2006, respectively. Scores from SAT I and SAT II tests carried out 2001 and the scores of 2006 along with high school grades were compared in regards to the students' undergraduate grade point average. The results showed that high school grades were a better predictor of students' undergraduate success. Moreover, when each component of the new and old SAT was examined, the second best predictor of undergraduate success was to be found new SAT writing test which was followed by the new SAT verbal test.

One of the drawbacks of the various previous studies is the fact that some other variables affecting the academic success were just ignored. Therefore, the study carried out by Olani (2008) has utmost importance as not only prior academic achievements such as GPAs, university entrance exam scores but also psychological

variables such as academic self-efficacy were taken into account. In order to find out the predictors of first year university students' academic success, the data were collected from 3301 students, 214 of whom filled in a self-report related to psychological variables. By using a multiple regression method, the degree to which the academic success of first year university students (GPAs) could be predicted from the two variables namely their prior academic achievements and psychological variables. The results taking the gender differences into consideration showed that while the combination of the two variables explained the 17% variance in the academic success of first year university students (GPAs), the psychological variables alone can be accounted for 4% variance in the students GPAs. The variance was found relatively higher in females than males, being 34% and 15%, respectively. This study also supports the findings of previous studies in the literature that students' high school GPAs are strong predictors of their GPAs at a university level.

2.3.2 Studies at Local Context

Ösken (1999) investigated the predictive validity of six midterms and a final assessment test conducted in the Department of Basic English at Hacettepe University. 3570 scores in total were analyzed in order to examine the predictive validity of these assessment practices. In addition, a questionnaire was given to 510 of Beginner level students and 34 instructors to have their opinions about the content of the course and effectiveness of the final assessment test. The analysis of the data by using Pearson Product Moment Correlation Coefficient showed that the predictive value of these assessment practices was very high (differing from r = .90 to r =.72). Meanwhile, the results of the questionnaire showed that participants think all the skills covered in the courses should be assessed in the final assessment test; therefore, she concluded that a need for a change is necessary.

Güzeller (2005) conducted a study with the 586 randomly selected students of 7th grade who took the OKOSYS exam (Student Selection and Placement tests carried

out by Ministry of National Education every year for Secondary School Students) in 2002. Both their GPAs and OKOSYS exam results were gathered and analyzed by using Correlational analysis. At the end of the study, a positive correlation was found, as 39% of the variance in their OKOSYS exam results can be accounted by their GPAs. Moreover, the better scores in each of the following courses, science, Turkish, social sciences resulted in higher OKOSYS exam results.

A similar study concerning higher education, more specifically the graduate achievement, was carried out in Ankara University in Turkey by Tokat and Demirtaşlı in 2004. In their study, while the independent variables were Graduate Education Entrance Examination (GEEE) used for admission to graduate programs with or without thesis, Undergraduate Point Average (UPA), Interview and Composite Evaluation Measure (CEM), the dependent variable was the students' graduate achievement. 251 graduate students participated in the study. For the data analysis, Pearson Product Moment Correlation and Multiple Regression Analysis was used and not only was the relationship between the dependent and independent variables investigated but also the one between each independent variable was also examined both for the thesis group and without thesis group. The results of this study show that while there was no significant relationship between the interview scores and CEM for the without thesis group, a statistically significant relationship was found for the general and the thesis group; however, it was not a strong one, which means there is not a strong relationship between these two variables. As for the GPA, GEEE and interview scores were a strong predictor for the general group.

A very similar study was conducted by Kan in 2005. However, this time, rather than having OKOSYS exam results as dependent variable, OSS scores was examined in relation with students' academic grades. 103 students studying at a private high school participated in the study. Their exam scores in the areas of literature, history, math, physics, chemistry and biology and their OSS scores were obtained and

analyzed. The highest correlation was found between the students' physics exam scores and their OSS physics scores.

Another similar study which examined the relationship between the two variables was carried out by Küçük (2007). However, the focus was not only on the predictive validity of the achievement tests but also on face validity of them. With these aims, two questionnaires were used to gather information about the face validity of the tests administered; one for the 29 teachers working at the Foreign Languages Preparatory School of Zonguldak Kara Elmas University and the other for the beginner level students of the same university. For predictive validity, the correlations between students' first term averages, second term averages, cumulative averages and the final assessment scores were investigated. The related data were obtained from 477 beginner level students. In addition, 52 of them filled in the questionnaire so that their perceptions related to face validity of the tests could be gathered. The results showed that the achievement tests have face validity to a high degree both according to the teachers and the students. To investigate the correlation between the variables, Pearson Product Moment Correlation was used and a significant positive correlation was found between the first term averages, second term averages, cumulative averages and the final assessment scores.

Karakaya (2007) examined the relationship between OSS scores and academic achievement, yet the participants were different. Rather than having high school students, the ones studying in the teaching programs in Faculty of Education were chosen. 691 students from five different universities were chosen and their OSS scores, high school grade point average (HSGPA) and freshman grade point average (FGPA) were obtained. After the analysis of the relationship between the variables, a positive correlation ranging from low to medium according to the department that the participants were studying at was found for all the variables. However, the higher

relationship was found between the scores taken from the students of chemistry and math department.

As her master thesis, in 2009, Sevindik also investigated the relationship between academic achievement score and scores of SBS exam of 2008. In her study, first of all, she gathered the grades of students taking the exam along with their Turkish, math, science, English and social studies grades. The participants of the study were 1034 in total, 525 from 6th and 519 from the 7th grade. The relationship between the two variables, exam scores and students' grade, was checked. At the end of her study, a positive correlation was found between all types of subjects and the exam scores but interpreted the correlation scores around .32 as strong even if they are regarded as moderate by most of the researchers and statisticians.

2.4 The Relationship between Nationwide and International Exams and Their Predictive Power

There are also various studies examining a nationwide exam and an international one, primarily TOEFL IBT (Test of English as a Foreign Language- Internet-based), whose scores can be used in a wide range of purposes in English-speaking countries. The study carried out in 2012 by Brown, Davis and Takahashi also aimed to find equivalencies between EIKEN (Test in Practical English Proficiency), and TOEFL IBT exam. The EIKEN English examinations are used in five countries as an admission requirement. Therefore, that they can be used to predict TOEFL IBT scores was the focus of attention in this study. 28.5% of the participants ($n = 35$) were male whereas 71.5% of them were female ($n = 88$), making a total of 123. While their ages ranged from 18 to 39, there were many differences in terms of their mother tongue. As the minimum number of years and the maximum that they were given formal English instruction were one year and 25 years, their TOEFL IBT scores also ranged from the lowest 47 to the highest 119. All the participants were assigned to an EIKEN grade level depending on their TOEFL IBT scores and the

EIKEN tests including reading, listening and writing were administered. For the analysis, one-way analysis of variance (ANOVA) is used as there is only one dependent variable which is the score of TOEFL IBT and one independent variable, their EIKEN test scores. The results showed a low to correlation among all of the subtests of EIKEN and TOEFL IBT.

2.5 The Relationship between International Exams and Academic Success and their Predictive Power

The study carried out in 1999 (AL-Ansari & Al-Musawi) investigated the relationship between two different international exams and academic success. This time, TOEFL and the first certificate of English (FCE) were independent variables and the academic success of students studying at the University of Bahrain (their GAPs) was dependent variable. The number of participants was 86. At the end of the regression analysis, FCE was found to be a better predictor of students' academic success compared to TOEFL. Especially the parts of FCE including cloze test and sentence transformation explained the highest variance in the GPAs of the students. Having such results led the researchers to share their ideas about the reconsideration of using TOEFL test as a test instrument in a second language learning context.

In order to investigate the relationship between TOEFL IBT scores and grade point average (GPA), a study carried out by Cho, and Bridgeman in 2012. They collected academic records including GPAs, TOEFL IBT, SAT and GRE (Graduate Record Examination) scores of 1850 graduate and 744 undergraduate students from 10 different institutions. According to the academic category, students, individual courses or majors were divided into four sub-groups namely, humanities and arts (HA), sciences and engineering (SE), social sciences (SS) and lastly business (BU) and the GPAs in these four fields were obtained. At the end of the correlation and multiple regression analysis, TOEFL IBT was found to explain 3% of the variance in

both graduate and undergraduate students of GPAs across disciplines. Similar results were also found for the relationship between SAT and GRE scores and GPAs.

2.6 The Relationship between Achievement and Other Variables

In this section, both the studies in local and global context examining other variables which may affect students' achievement were presented. Roby (2003) examined the relationship between attendance and students' achievement. For this aim, the results of Ohio Proficiency tests taken by the students of grades four, six, nine and twelve in 3.171 schools and the attendance averages of the same students were obtained. In order to find out whether there is a positive correlation between the variables or not, Pearson's correlation statistic was used. While the results showed a medium positive correlation between the achievements of students in grade four, six and twelve and their attendance, a strong positive correlation was found in the ninth grade.

The other variable taken into account by various researchers is the time that students spend for homework. In 2007, McMullen examined the impact of increased amount of homework on students' academic success. When previous literature was revised, it was concluded that almost all of the studies showed evidence saying that there is a positive relationship between these two variables. For the data collection, 7902 students were asked to fill in a survey including an achievement test in four different areas; math, science, English and history. There were also some other questions related to time spent on homework. At the end of the study, similar to previous studies, a positive relationship was found.

İleri (2013) also investigated the relationship between students' homework grades and TOEFL ITP in the Department of Foreign Languages at TOBB ETU. She gave a questionnaire to 286 students and 23 instructors in order to get their opinions towards homework. At the end of the data analysis, she found out that even though the percentage of homework done by the students is influential for their success in the

assessment practices conducted throughout the term, no significant relationship is found between homework and the students' scores of TOEFL ITP.

As in the case of most of the previous studies, only two scores or two variables are correlated in order to find out the relationship between them. However, as it will not be enough only to check the types of relationship between the variables, a deeper analysis was planned to be carried out in this study. Therefore, the predictive power of the assessment practices to TOEFL ITP has utmost importance. Moreover, in the institution, TOBB ETU, neither the students' nor the teachers' opinion related to assessment practices carried out by testing and curriculum unit have been investigated before. Therefore, conducting the present study will fill this gap.

CHAPTER III

METHOD of DATA COLLECTION and ANALYSIS

This present study focuses on the relationship between various assessment practices carried out at TOBB ETU and TOEFL ITP. It also attempts to investigate the predictive power of these tests to TOEFL ITP. Moreover, the perception of both the students and instructors related to the predictive power of the assessment practices and TOEFL ITP is also a topic of interest. Therefore, following research questions are addressed:

1. How well do the different types of achievement tests including midterms, quizzes, readers and placement exam results predict the TOEFL ITP scores of students studying at the preparatory program at TOBB University?
2. Is there a significantly positive relationship among the different types of achievement tests including midterms, quizzes, readers and placement exam grades and TOEFL ITP scores of students studying at the preparatory program at TOBB University?
3. What are the students' and instructors' perceptions towards the relationship between all these assessment practices and TOEFL ITP?

This chapter covers and describes the overall design of the study, participants, data collection instruments, reliability and validity of the instruments and data analysis procedure.

3.1 Overall Design of the Study

The main aim of this study is to analyze the predictive power of various assessment practices carried out in the Department of Foreign Languages at TOBB ETU by using Pearson Product Moment Correlation. As suggested by Aiken in 1997, the predictive design is intended for forecasting or predicting specific future outcome. Therefore, firstly, the relationship between all the variables and TOEFL ITP was examined for each level of Department of Foreign Languages. Moreover, the extent that the combination of the variables predicts students' scores of TOEFL ITP was investigated for each level separately. In addition, both the students and instructors filled in a questionnaire to share their opinions about the effectiveness of these assessment practices. The data regarding the scores of various assessment tests were obtained from the student affairs information system with the consent of the administration of TOBB ETU. Besides, to obtain the information related to students' and teachers' perception, a questionnaire was formed and utilized.

As it can also be understood from the Figure 4 above, the relationship between the assessment practices carried out by the Testing and Evaluation Unit and TOEFL ITP scores along with students' and instructors' opinion on the topic is the primary concern of this study.

```
                    ┌─────────────────────────┐
                    │    Design of the Study   │
                    └─────────────────────────┘
```

| Quantitative Data | Qualitative Data |

| The predictive validity of various assessment tests carried out by Testing and Evaluation Unit, TOBB ETU | Both the instructors' and students' opinion on the predictive power of assessment practices to TOEFL ITP scores |

Placement Exam Results

Midterms (Each one separately)
Quizzes (Each one separately)
Readers (The average of all)
&
TOEFL ITP

Instructors' Questionnaire
&
Students' Questionnaire

Figure 4: The Design of the Study and the Illustration of the Variables

3.2 Characteristics of the Participants

The study was conducted with the data of the students of Department of Foreign Languages at TOBB ETU, in the 2012-2013 academic year. Both the students who fail because of absenteeism and the ones who take a different exam such as FCE, TOEFL IBT or IELTS rather than institutional TOEFL ITP to start studying in their departments were not taken into account.

There are two requirements for the students to take TOEFL ITP in order to continue their education in their departments. The first one is not to go beyond the absence limit as they are required to be in class 90% of each term. The next one is to gather a total of 65 or more from the assessment methods used in order to take TOEFL ITP at the beginning of July, which is the end of the year in the preparatory program. However, this situation differs for upper-intermediate (C Level) students. They are

given a chance to take TOEFL at the end of December, if they fulfill the same requirements. For this study, all the students who meet these two requirements were taken into account. A more detailed illustration of the participants is provided in Figure 5 below.
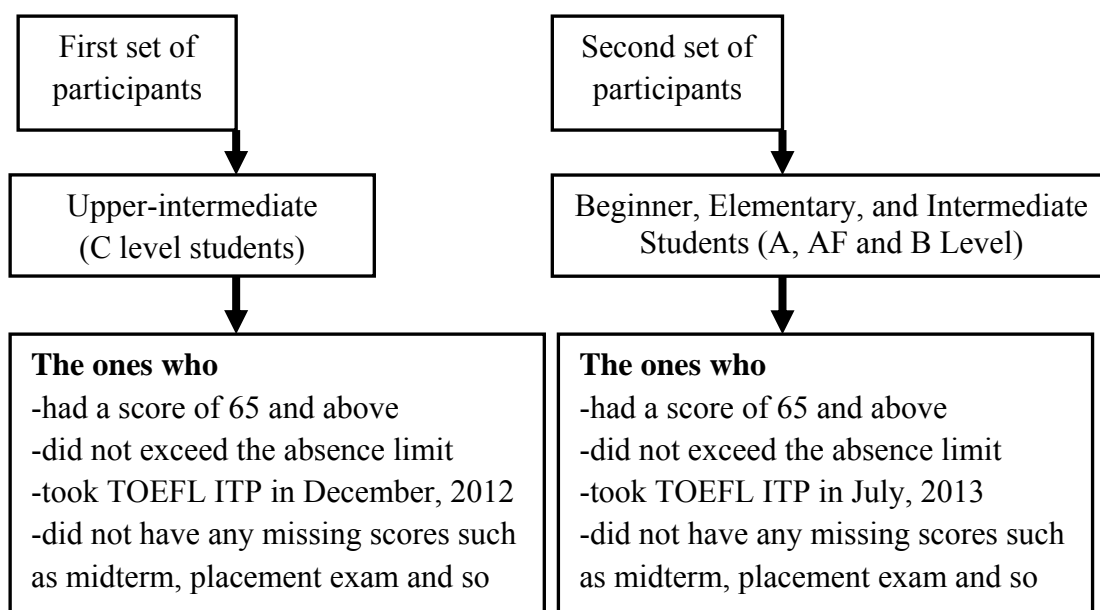
Figure 5: Detailed Illustration of the Participants

For the questionnaire, there are two groups of participants. As it can be understood from the table below, there are 421 students filling the questionnaire in total. As for the instructors out of 53 who teach in the program, 48 of them filled in the questionnaire.

Table 1: *The Number and Percentage of Students Given the Questionnaire*

Level	Total number of students	Total number of students filling the questionnaire
A	152	113
B	155	128
AF	386	180

As illustrated in the chart above, while the number of AF Level students filling the questionnaire is the highest, the number of A Level students was the lowest (See Table 1).

3.2.1. Instructors

In order to gather information related to the demographic data of the participants, a questionnaire was given to all instructors working at TOBB ETU during the 2012-2013 academic year. However, out of 53 instructors in total, 48 returned the questionnaire ($p = 90.6\%$). Therefore, the study involved 48 instructors. While 47 of the participants were female ($p = 97.9\%$), only one of them were male ($p = 2.1$). Their age average was 27 ranging between 22 to 38. The first question in the questionnaire was about their experience in teaching English (See Table 2). Out of 48, four of them ($p = 8.3\%$) said that they had been teaching less than a year. While 16 ($p = 33.3\%$) had one to three years of experience, 20 of them ($p = 41.7\%$) had three to six years of experience. Lastly, eight instructors ($p = 16.7\%$) told that they had been teaching English for more than six years.

Table 2: *The Results of Instructors' Questionnaire: Question 1 - Years of Experience*

	N	*%*
Less than a year	4	8.3
One to three years	16	33.3
Three to six years	20	41.7
More than six years	8	16.7

The next question was about the length of time that they have been working at TOBB ETU. Twenty participants (p = 41.7%) had been working for this institution for less than a year, 18 of them (p = 37.5%) had been working for one to three years, six of them (p = 12.5%) had been working at this institution between three to six years and lastly, four of them (p = 8.3%) had been working for more than six years (See Table 3).

Table 3: *The Results of Instructors' Questionnaire: Question 2 - Years of Experience at TOBB ETU*

	N	%
Less than a year	20	41.7
One to three years	18	37.5
Three to six years	6	12.5
More than six years	4	8.3

When we look at the number of instructors who have been in a foreign country where English is the official language, 13 (p = 27.1%) stated that they had never been to an English-speaking country. Four (p = 8.3%) of them stayed there for less than a week, 16 (p = 33.3%) of them stayed there from a week to a month, 13 (p = 27.1%) of them stayed there up to one year and lastly, two instructors (p = 4.2%) stayed there for more than a year (See Table 4).

Table 4: *The Results of Instructors' Questionnaire: Question 5 and 6 - Duration of Staying in an English-Speaking Country*

	N	%
No	13	27.1
Less than a week	4	8.3
From a week to a month	16	33.3
Up to one year	13	27.1
More than a year	2	4.2

3.2.2. Students

The first part of the questionnaire distributed to the students is designed to gather demographic data. The study involved 421 students with the largest group being the AF Level students ($n = 180$) (See Table 5 below). While 208 were male (p = 49.4%), 213 (p = 50.6%) were female. Their age ranges from 18 to 34 with an average of 19.

Table 5: *The Number of Students for Each Level Filling the Questionnaire*

		N	*%*	*Cum. Percent*
Valid	B Level	128	30.4	30.4
	A Level	113	26.8	57.2
	AF Level	180	42.8	100.0
	Total	421	100.0	

In Figure 6 below, the number of students and their department are illustrated. While the students studying at the Department of Electrical and Electronics Engineering ($n = 47$) are the largest in number, the number of students studying at the English Language and Literature Department is the lowest ($n = 2$).
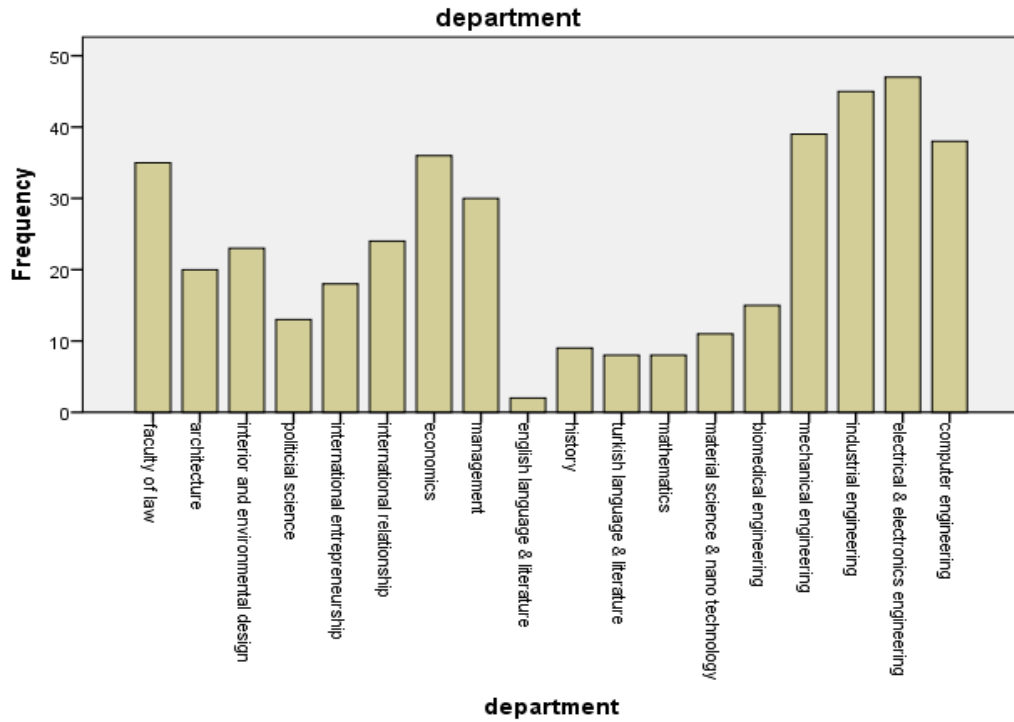
Figure 6: Distribution of Students According to their Departments

The next question in the questionnaire was about their experience in learning English and the type of school that they graduated from. Out of 421, 243 (*p* = 57.7%) said that they had been learning English for more than six years. While 81 students (*p* = 19.2%) had less than one year of study for English, 65 (*p* = 15.4%) had three to six years of study. Lastly, 32 students (*p* = 7.6%) told that they had been learning English from one year to three years (See Table 6).

Table 6: *The Results of the Students' Questionnaire: Question 4 - Experience of Learning English*

		N	%	Cum.Percent
	Less than one year	81	19.2	19.2
	1-3 years	32	7.6	26.8
Valid	3-6 years	65	15.4	42.3
	More than 6 years	243	57.7	100.0
	Total	421	100.0	

As it can be understood from the table below, 225 students (*p* = 53.4%) graduated from an Anatolian high school. In addition, there are only four students who graduated from a vocational high school. There are four students who go for the option "other". While three of them graduated from a social sciences high school, one of them graduated from a religious vocational high school (See Table 7).

Table 7: *The Results of the Students' Questionnaire: Question 5 - Types of Graduate Schools*

		N	%	Cum Percent
	College	66	15.7	15.7
	Regular High School	66	15.7	31.4
	Vocational High School	4	1.0	32.3
Valid	Anatolian High School	225	53.4	85.7
	Science High School	32	7.6	93.3
	Anatolian Teacher Training High S.	24	5.7	99.0
	Others	4	1.0	100.0
	Total	421	100.0	

The students were also asked to indicate whether they had ever been an English-speaking country before, whether they have any friends or relatives who are native speakers of English and whether they had ever taken a course from a native teacher before. Three hundred and sixteen students (*p* =75.1%) had not been to an English speaking country before. While 54 students (*p* = 12.8%) stayed in an English-

speaking country for a week to a month, 43 of them (p = 10.2%) stayed in one of these countries less than a week. Out of 421, only seven of them stayed in an English-speaking country up to one year and only one of them spent more than a year there (See Table 8).

Table 8: *The Results of the Students' Questionnaire: Question 6 and 7 - Duration of Staying in an English-Speaking Country*

		N	%	Cum Percent
	Yes (Less than a week)	43	10.2	10.2
	Yes (A week to a month)	54	12.8	23.0
Valid	Yes (Up to a year)	7	1.7	24.7
	Yes (More than a year)	1	.2	24.9
	No	316	75.1	100.0
	Total	421	100.0	

For the next question, while 326 students (p = 77.4%) said that they had no friends or relatives speaking English as a native speaker, 95 (p = 22.6%) said that they had. For the ones who had English native speaker friends or relatives, two follow-up questions were asked: how often they get in touch with them and what kinds of communication tools are used. Most of them preferred to use social networks or text-messaging, yet there were a few who talked with their native relatives on the phone and less frequently face to face (See Table 9).

Table 9: *The Results of the Students' Questionnaire: Question 8 - Having a Friends or Relatives Speaking English as a Native Speaker*

		N	%	Cum. Percent
	yes	95	22.6	22.6
Valid	No	326	77.4	100.0
	Total	421	100.0	

Three hundred and five of the students (p = 72.4%) said that they had never had a native speaker teacher of English before while 116 (p = 27.6%) said that they had been taught by a native speaker. As a follow-up question, they were asked to indicate how long they had been taught by this teacher; their responses varied a month to three years. They were also asked how often they had been taught by the native speaker teacher of English. Their replies also changed from an hour to ten hours per week (See Table 10).

Table 10: *The Results of the Students' Questionnaire: Question 11 - Having a Native Speaker Teacher of English or Not*

		N	%	Cum. Percent
	Yes	116	27.6	27.6
Valid	No	305	72.4	100.0
	Total	421	100.0	

Lastly, the students were asked two open-ended questions. The first question was how many hours a day they study English and the second one was what kind of activities they are engaged in order to improve their English. There was a great variation in their answers to both questions. As for the first one, there were some students who study only 15 minutes per day, and some others who spend more than three hours. For the last question, the students listed different types of activities such as watching T.V series or movies with or without subtitles, solving problems, reading books, playing online games, listening to the news, benefitting from some websites-ted.com, history.com- or different applications for phones.

3.3 Data Collection Instruments

In this study, students test scores were obtained from the uniform system with the consent of the administration. For the data which aim to discover both the students' and instructors' opinion, two different questionnaires were developed by the researcher herself and given to those groups. While developing the data collection

tools, some principles suggested by Slavin (2007) are followed. Therefore, the questions were simple and clear, double negatives were avoided, and only important and relevant questions were asked. After getting feedback from various instructors, the final version was formed.

In the first part of the questionnaire for the instructors, some demographic questions were asked about their age, gender, years of experience at TOBB ETU and years of experience overall, and whether they had been in a foreign country where English was spoken and how long they had been there. This part included six items. In the questionnaire for the students, in addition to age and gender, their department, graduate high school, the presence of any native speaker of English among the family members or friends, types of communication tools that they used to interact with them, time spent to improve their English level and lastly kinds of activities that they dealt with in order to advance their English were asked. There were 14 questions in this part.

As the focus of the second part was on the opinions of students' about assessment practices carried out at TOBB ETU, relevant questions were asked including the effectiveness of these practices on their TOEFL ITP scores. There were three open-ended questions along with a Likert-scale which required them to mark a point from one to six to determine the effectiveness of these assessment practices. Instead of having a five-point Likert scale where more of the participants go for number three, the middle one, which does not give any specific information about their choice, a six-point-scale was used where there is a clear-cut distinction between the points. After the questionnaire was designed, it was shown to the instructors working at the testing and curriculum unit of the institutions for feedback.

3.4 Validity and Reliability of the Instruments

Validity is one of the key factors that need to be kept in mind while designing or preparing an instrument. All three types of validity -content, criterion and construct-, are considered in the process of creating a new instrument, a questionnaire in this case for the present study. Content validity, as the name suggests, refers to the relevance of the questionnaire to what is intended to be covered. For this study, three instructors working in the Testing and Evaluation Unit, and five more from the Curriculum and Development Unit were asked to share their opinions about the content of both students' and instructors' questionnaire. During that process, they filled in the questionnaire to point out the ambiguous parts and overall appearance of the questions. Moreover, one more instructor working in the Testing Unit at the Department of Modern Languages at Middle East Technical University shared her ideas about the content of the questionnaire and gave feedback. After the feedback, some more questions were added, ambiguous parts were changed and lastly the format of the questionnaires was changed.

There are some threats to validity in survey research such as mortality and instrumentation threat (Frankel & Wallen, 2006). To begin with, mortality threat which means loss of participants during the data collection process was not a problem in the study. Thanks to detailed planning, data collection was completed at the same time for all the students and within three days for the instructors. Moreover, in order to avoid instrumentation threat, the questionnaire was prepared in such a way that participants would not get exhausted or unwilling to complete it.

A small sample size in studies might be a problem while investigating the relationship between the variables or examining the effect of one to the other. However, in this study, the number of participants was adequate enough to have power to make some conclusions and find answers to the research questions.

The questionnaire was implemented in the native language of the participants in order to make them express themselves in a more comfortable way. After collecting the data instruments, the comments were translated into English by the researcher herself. For the translation reliability, some problematic parts were shown to one of the instructors who was also a graduate of an Foreign Language Education department of METU.

3.5 Data Analysis Procedure

For this study, midterm, quiz and reader grades of students from all levels were obtained. While midterm and quiz results were entered to SPSS, for the reader results, the average was taken and the obtained score was written there. As for the questionnaire (see Appendix A, B, C and D) which not only covers questions related to demographic data but also open-ended questions related to assessment practices, thematic analysis was used to find and examine patterns within the questionnaire. In other words, the results of the questionnaire were analyzed to gain better understanding of the students' and instructors' opinion after finding the common points based on their answers. As suggested by Yıldırım and Şimşek (1999), in thematic analysis, there are some steps. To begin with, first the data are read and grouped into different categories. These categorization can be done either before or during the analysis. After that, with quotations from the responses of the participants, the findings and the implications are presented to the readers.

First of all, all the results were entered to SPSS in order to see their correlation one by one. Later on, in order to find answers to both research questions, all the variables were correlated and graphs were drawn with SPSS output by using Pearson Product Moment Correlation and Multiple Regression as all the variables are interval measures.

In order to analyze the relationship between the variables, the Pearson's correlation statistics was utilized. Pearson's r is established for the following:

1. AF (Beginner) level midterms, quizzes, readers, placement exam results, and their TOEFL ITP scores

2. A (Elementary) level midterms, quizzes, readers, placement exam results, and their TOEFL ITP scores

3. B (Intermediate) level midterms, quizzes, readers, placement exam results, and their TOEFL ITP scores

4. C (Upper-Intermediate) level midterms, quizzes, readers, placement exam results, and their TOEFL ITP scores

A detailed figure is provided below to show the types of data gathered and the sources of that data.

Data Set 1	Data Set 2	Data Set 3
Types of data - Placement exam results - Midterms - Quizzes - Readers	**Types of data** - Both the students and teachers' perception on the relationship between assessment practices and TOEFL ITP	**Types of data** -Scores of TOEFL ITP
Sources of data - Students Affairs Information system of TOBB ETU	**Sources of data** - Questionnaire distributed to the students and instructors	**Sources of data** - ETS sent the results to the institution

Figure 7: Types of Data and the Sources Used to Gather Them

For the first research question, (how well do the different types of achievement tests including midterms, quizzes, reader and placement exam grades predict the TOEFL ITP scores of students studying at the preparatory program at TOBB University?) the

scores of the achievement tests were obtained and their relationship along with their predictive power to the scores of TOEFL ITP was analyzed using multiple regression and correlation analysis.

For the second research question, (is there a significantly positive relationship among the different types of achievement tests including midterms, quizzes, readers and placement exam grades and TOEFL ITP scores of students studying at the preparatory program at TOBB University?), the data were collected and by the help of the same statistical analysis, the graphs and the degree of prediction were drawn and understood, respectively.

For the third research question, (what are the students and instructors perceptions towards the relationship between all these assessment practices and TOEFL ITP?), two different questionnaires were formed and given to two different groups; students and instructors. Common points were placed under various categories. By the help of this coding, main issues raised by the students and instructors were understood.

CHAPTER IV

RESULTS of the DATA ANALYSIS

In this chapter, the results of the qualitative data and the quantitative data are presented. In Figure 8 below, a detailed illustration of the way this section is presented can be found.

Results of Quantitative Data	Results of Qualitative Data
✓ C Level Analysis ✓ AF Level Analysis ✓ A Level Analysis ✓ B Level Analysis	✓ Results of the Instructor Questionnaire ✓ Results of the Student Questionnaire ✓ Comparison of the Results of Instructor and Student Questionnaire
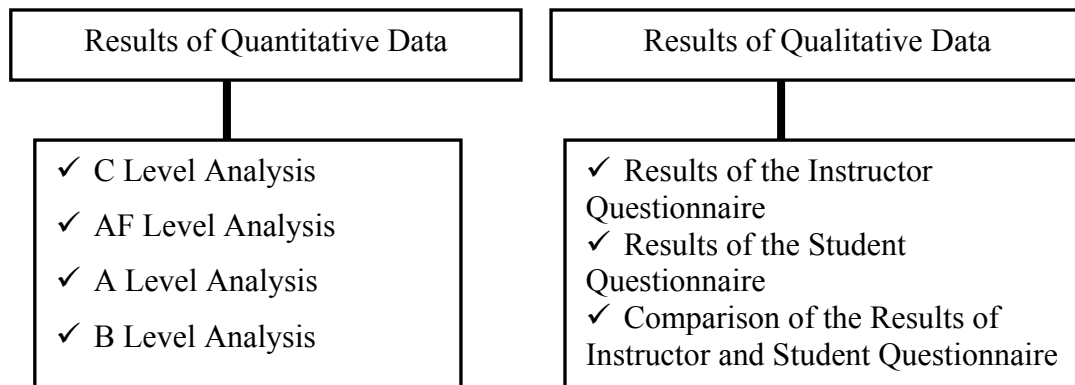
Figure 8: Detailed Presentation of Data Analysis and Result Section

4.1. Results of the Quantitative Data

Pearson Correlation and Multiple Regression were used to figure out the relationship between midterms, quizzes, reader grades and placement exam results and TOEFL ITP carried out at the end of the semester. In this part, the results of the quantitative data are presented for each level separately beginning with C Level.

4.1.1. C Level (Upper-Intermediate Students)

At the Department of Foreign Languages, TOBB ETU, C Level refers to upper-intermediate students who are expected to take TOEFL test at the end of December. As it can be seen in Table 11 below, the descriptive analysis of the data in relation to all the assessment practices for C Level presents that among the quizzes while Quiz-1 (CQI) has the lowest mean *(M= 57.02, SD= 13.87),* Quiz-2 (CQII) has the highest mean *(M= 78.70, SD= 14.46).* Among the midterms, Midterm-1 (CMI) has the lowest mean *(M= 59.19, SD= 10.54)* and Midterm-2 (CMII) has the highest mean *(M= 73.04, SD= 10.79).* While students' placement exam grades (CPLACE) range between 9.60 and 98.40, with an average of 64.28, their TOEFL scores (CTOEFL) range between 433 and 543, with an average of 474.70 (See Table 11).

Table 11: *Descriptive Statistics for C Level*

	N	*Minimum*	*Maximum*	*M*	*SD*
CMI	**150**	**34.10**	**81.70**	**59.19**	**10.54**
CMII	**150**	**45.00**	**96.40**	**73.04**	**10.79**
CMIII	150	11.55	92.24	66.15	11.79
CQI	**150**	**22.50**	**92.50**	**57.02**	**13.87**
CQII	**150**	**44.00**	**100.00**	**78.70**	**14.46**
CQIII	150	32.50	96.25	70.88	13.35
CRTOTAL	150	22.67	100.00	78.38	15.38
CPLACE	**150**	**9.60**	**98.40**	**64.28**	**23.91**
CTOEFL	150	433.00	543.00	474.70	20.38
Valid N (listwise)	150				

To see the relationship between the variables one by one, Pearson Correlation was used. However, as the main focus of this study is TOEFL scores, only the relationship between all other variables and TOEFL scores is analyzed. First of all, the relationship between all the quizzes, midterms and the students' scores of TOEFL ITP is checked. As it can be seen from Table 12 below, there is a significant relationship between all the quizzes and TOEFL, as well as between all the midterms and TOEFL scores. While the highest significant correlation is found between Quiz-

1 (CQI) and TOEFL, $r = .49$ $n = 150$, $p > .05$, two-tailed test, the lowest but again significant correlation is found between Quiz-2 (CQII) and TOEFL, $r = .30$, $n = 150$, $p > .05$, two-tailed test. Likewise, while the highest significant correlation is found between Midterm-1 (CMTI) and TOEFL, $r = .49$ $n = 150$, $p > .05$, two-tailed test, the lowest but again significant correlation is found between Midterm-3 (CMTIII) and TOEFL, $r = .40$, $n = 150$, $p > .05$, two-tailed test.

In Table 12 below, correlation between quizzes, midterms and TOEFL scores for C Level can be found:

Table 12: *Correlations for C Level Midterms, Quizzes and TOEFL ITP*

		CMTI	CMTII	CMTIII	CQI	CQII	CQIII	CTOEFL
CMTI	Pearson Correlation		.77**	.67**	.70**	.45**	.60**	**.49****
	Sig. (2-tailed)		.00	.00	.00	.00	.00	**.00**
	N		150	150	150	150	150	**150**
CMTII	Pearson Correlation			.65**	.66**	.54**	.66**	**.44****
	Sig. (2-tailed)			.00	.000	.00	.00	**.00**
	N			150	150	150	150	**150**
CMTIII	Pearson Correlation				.53**	.40**	.57**	**.40****
	Sig. (2-tailed)				.00	.00	.00	**.00**
	N				150	150	150	**150**
CQI	Pearson Correlation					.34**	.52**	**.49****
	Sig. (2-tailed)					.00	.00	**.00**
	N					150	150	**150**
CQII	Pearson Correlation						.32**	**.30****
	Sig. (2-tailed)						.00	**.00**
	N						150	**150**
CQIII	Pearson Correlation							**.37****
	Sig. (2-tailed)							**.00**
	N							**150**

**. Correlation is significant at the 0.01 level (2-tailed).

As a second step, the relationship between the average of readers, placement exam results and the students' scores of TOEFL ITP is checked. As it can be seen from Table 13 below, there is a significant relationship only between the placement exam results of the students and TOEFL, $r = .27$ $n = 150$, $p > .05$, two-tailed test. The correlation between these three variables and TOEFL scores for C Level can be found below:

Table 13: *Correlations for C Level Reader Exam, Placement Exam and TOEFL ITP*

		CRTOTAL	CPLACE	CTOEFL
	Pearson Correlation		.21*	**.15**
CRTOTAL	Sig. (2-tailed)		.01	**.06**
	N		150	**150**
	Pearson Correlation			**.27****
CPLACE	Sig. (2-tailed)			**.00**
	N			**150**

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

In the figures below, for the three highest correlations (the one between Quiz-1 (CQI) and TOEFL, Midterm-1 (CMI) and TOEFL, and placement exam results and TOEFL), scatter plots are drawn. As it can be seen in all three figures, there appears a positive strong linear relationship between these variables.
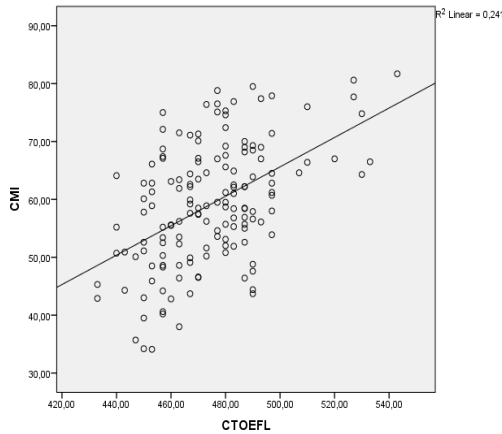
Figure 9: Scatter Plot of Students'
Quiz-1 Results and TOEFL
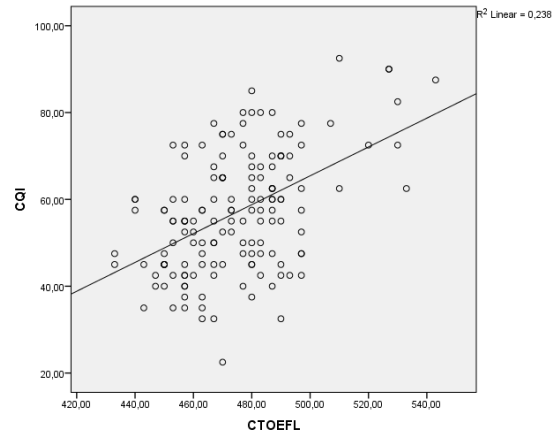
Figure 10: Scatter Plot of Students'
Midterm-1 results and TOEFL
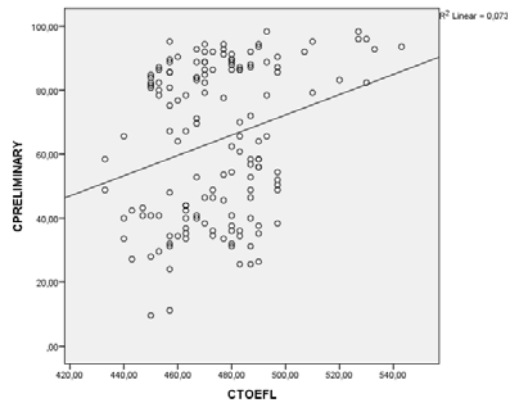
Figure 11: Scatter Plot of Students' Placement Exam Results and TOEFL

In addition to checking the relationship between the variables and TOEFL ITP scores one by one, the relationship between the combination of all the variables and TOEFL ITP is checked for C Level. In Table 14 below, the magnitude of correlation coefficient (.31) indicates that students' TOEFL ITP scores are moderately related to the all the variables together. More specifically, it can be said that these variables can explain 31% of the variation in TOEFL scores of the students.

Table 14: *Model Summary for C Level*

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.56[a]	.31	.27	17.38

a. Predictors: (Constant). CPLACE. CQII. CQIII. CQI. CRTOTAL. CMIII. CMI. CMII

As it can be seen in Table 15 below, there is a positive significant relationship between the variables and TOEFL ITP scores of students $F_{(8,141)} = 7.98$, $p = 0.00$.

Table 15: *ANOVA Table for C Level*

Model		SS	Df	Mean Square	F	Sig.
	Regression	19285.81	8	2410.73	7.98	.00[b]
1	Residual	42573.69	141	301.94		
	Total	61859.50	149			

a. Dependent Variable: CTOEFL

b. Predictors: (Constant). CPLACE. CRTOTAL. CQIII. CQII. CQI. CTMIII. CTMI. CTMII

4.2.2 AF Level (Beginner Level Students)

AF Level refers to beginner level students who are expected to take TOEFL test at the beginning of July after receiving one year of English education. Their English level was the lowest among the other three levels. As seen in Table 17, the descriptive analysis of the data in relation to all assessment practices for AF Level presents that among the quizzes while Quiz-1 (AFQI) has the highest mean *(M= 94.48, SD= 5.41),* Quiz-9 (AFQIX) has the lowest mean *(M= 71.13, SD= 13.09).* Among the midterms, Midterm-9 (AFMTIX) has the lowest mean *(M= 62.53, SD= 10.97)* and Midterm-5 (AFMTV) has the highest mean *(M= 78.52, SD= 8.93).* In addition, while students' placement exam grades (AFPLACE) range between 7.20 and 48.80, with an average of 38.89, their TOEFL scores (AFTOEFL) range between 420 and 587, with an average of 481.52.

Table 16: *Descriptive Statistics for AF Level*

	N	Minimum	Maximum	M	SD
AFPLACE	**101**	**7.20**	**48.80**	**38.89**	**9.27**
AFQI	**101**	**70.10**	**100.00**	**94.48**	**5.41**
AFQII	101	49.50	100.00	81.92	12.37
AFQIII	101	58.50	96.75	79.18	8.39
AFQIV	101	44.00	100.00	88.21	10.60
AFQV	101	52.50	97.50	78.29	9.74
AFQVI	101	58.75	98.75	79.94	9.80
AFQVII	101	53.88	99.25	80.89	8.90
AFQVIII	101	28.50	100.00	81.98	13.48
AFQIX	**101**	**40.00**	**97.50**	**71.13**	**13.09**
AFQX	101	35.00	97.50	72.15	11.81
AFQXI	101	56.00	96.38	76.19	9.65
AFQXII	101	51.00	100.00	83.25	11.19
AFMTI	101	48.75	97.50	78.22	9.44
AFMTII	101	54.20	96.40	77.55	8.36
AFMTIII	101	45.00	96.40	77.97	10.42
AFMTIV	101	54.50	96.00	77.46	9.13
AFMTV	**101**	**55.50**	**96.00**	**78.52**	**8.93**
AFMTVI	101	52.50	99.00	75.02	10.71
AFMTVII	101	41.00	93.40	68.31	11.12
AFMTVIII	101	48.25	90.96	64.28	8.37
AFMTIX	**101**	**39.95**	**94.03**	**62.53**	**10.97**
AFMTX	101	32.51	92.71	63.61	10.99
AFRTOTAL	101	49.33	97.56	79.81	8.14
AFTOEFL	**101**	**420.00**	**587.00**	**481.52**	**30.25**
Valid N (listwise)	101				

To see the relationship between the variables one by one, Pearson Correlation was used. However, as the main focus of this study is TOEFL scores, only the relationship between all other variables and TOEFL scores is analyzed. First of all, the relationship between all the quizzes and the students' scores of TOEFL ITP is checked. As it can be seen from Table 17, there is a significant relationship between all the quizzes except Quiz-1 (AFQI) and TOEFL. While the highest significant correlation is found between Quiz-9 (AFQIX) and TOEFL, $r = .64$ $n = 101$, $p > .05$,

two-tailed test, the lowest but again significant correlation is found between Quiz-4 (AFQIV) and TOEFL, $r = 25$, $n = 101$, $p > .05$, two-tailed test. In table 17 on the next page, correlation between quizzes and TOEFL scores for AF Level can be found.

Secondly, the relationship between all the midterms and the students' scores of TOEFL ITP is checked. As it can be seen from Table 18 below, there is a significant relationship between all the midterms and TOEFL. While the highest significant correlation is found between Midterm-8 (AFMTVIII) and TOEFL, $r = .77$ $n = 101$, $p > .05$, two-tailed test, the lowest but again significant correlation is found between Midterm-1 (AFMTI) and TOEFL, $r = 41$, $n = 101$, $p > .05$, two-tailed test.

In table 18, correlation between midterms and TOEFL scores for AF Level can be found.

Table 17: *Correlations for AF Level Quizzes and TOEFL ITP*

		AFQI	AFQII	AFQIII	AFQIV	AFQV	AFQVI	AFQVII	AFQVIII	AFQIX	AFQX	AFQXI	AFQXII	AFTOEFL
AFQI	Pearson Correlation		.52**	.35**	.23*	.37**	.27**	.25*	.08	.15	.40**	.08	.19	.18
	Sig. (2-tailed)		.00	.00	.02	.00	.01	.01	.45	.14	.00	.41	.06	.08
	N		101	101	101	101	101	101	101	101	101	101	101	101
AFQII	Pearson Correlation			.47**	.45**	.54**	.57**	.51**	.21*	.41**	.43**	.34**	.36**	.34**
	Sig. (2-tailed)			.00	.00	.00	.00	.00	.04	.00	.00	.00	.00	.00
	N			101	101	101	101	101	101	101	101	101	101	101
AFQIII	Pearson Correlation				.41**	.51**	.42**	.49**	.34**	.41**	.30**	.44**	.26**	.37**
	Sig. (2-tailed)				.00	.00	.00	.00	.00	.00	.00	.00	.01	.00
	N				101	101	101	101	101	101	101	101	101	101
AFQIV	Pearson Correlation					.40**	.46**	.39**	.36**	.22*	.30**	.35**	.38**	.25*
	Sig. (2-tailed)					.00	.00	.00	.00	.03	.00	.00	.00	.01
	N					101	101	101	101	101	101	101	101	101
AFQV	Pearson Correlation						.61**	.44**	.34*	.61**	.52**	.48*	.33**	.54**
	Sig. (2-tailed)						.00	.000	.00	.00	.00	.00	.00	.00
	N						101	101	101	101	101	101	101	101
AFQVI	Pearson Correlation							.68**	.31**	.54**	.54**	.53**	.36**	.48**
	Sig. (2-tailed)							.00	.00	.00	.00	.00	.00	.00
	N							101	101	101	101	101	101	101
AFQVII	Pearson Correlation								.35**	.56**	.46**	.48**	.31**	.53**
	Sig. (2-tailed)								.00	.00	.00	.00	.00	.00
	N								101	101	101	101	101	101
AFQVIII	Pearson Correlation									.38**	.29**	.44**	.43**	.38**
	Sig. (2-tailed)									.00	.00	.00	.00	.00
	N									101	101	101	101	101
AFQIX	Pearson Correlation										.55**	.54**	.29**	.64**
	Sig. (2-tailed)										.00	.00	.00	.00
	N										101	101	101	101
AFQX	Pearson Correlation											.48**	.28**	.44**
	Sig. (2-tailed)											.00	.01	.00
	N											101	101	101
AFQXI	Pearson Correlation												.22*	.50**
	Sig. (2-tailed)												.03	.00
	N												101	101
AFQXII	Pearson Correlation													.42**
	Sig. (2-tailed)													.00
	N													101
AFTOEFL	Pearson Correlation													
	Sig. (2-tailed)													
	N													

Table 18: *Correlations for AF Level Midterms and TOEFL ITP*

		AFMTI	AFMTII	AFMTIII	AFMTIV	AFMTV	AFMTVI	AFMTVII	AFMTVIII	AFMTIX	AFMTX	AFTOEFL
AFMTI	Pearson Correlation		.78**	.60**	.64**	.52**	.47**	.19	.51**	.40**	.25*	.41**
	Sig. (2-tailed)		.00	.00	.00	.00	.00	.05	.00	.00	.01	.00
	N		101	101	101	101	101	101	101	101	101	101
AFMTII	Pearson Correlation			.68**	.67**	.59**	.58**	.42**	.62**	.48**	.35**	.50**
	Sig. (2-tailed)			.00	.00	.00	.00	.00	.00	.00	.00	.000
	N			101	101	101	101	101	101	101	101	101
AFMTIII	Pearson Correlation				.72**	.65**	.73**	.48**	.59**	.49**	.40**	.49**
	Sig. (2-tailed)				.00	.00	.00	.00	.00	.00	.00	.000
	N				101	101	101	101	101	101	101	101
AFMTIV	Pearson Correlation					.77**	.79**	.56**	.71**	.66**	.61**	.61**
	Sig. (2-tailed)					.00	.00	.00	.00	.00	.00	.000
	N					101	101	101	101	101	101	101
AFMTV	Pearson Correlation						.80**	.60**	.68**	.67**	.53**	.62**
	Sig. (2-tailed)						.00	.00	.00	.00	.00	.000
	N						101	101	101	101	101	101
AFMTVI	Pearson Correlation							.71**	.73**	.72**	.61**	.68**
	Sig. (2-tailed)							.00	.00	.00	.00	.000
	N							101	101	101	101	101
AFMTVII	Pearson Correlation								.64**	.65**	.63**	.62**
	Sig. (2-tailed)								.00	.00	.00	.000
	N								101	101	101	101
AFMTVIII	Pearson Correlation									.79**	.69**	.77**
	Sig. (2-tailed)									.00	.00	.000
	N									101	101	101
AFMTIX	Pearson Correlation										.69**	.72**
	Sig. (2-tailed)										.00	.000
	N										101	101
AFMTX	Pearson Correlation											.65**
	Sig. (2-tailed)											.000
	N											101
AFTOEFL	Pearson Correlation											
	Sig. (2-tailed)											
	N											

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

As for the last step, the relationship between the average of readers, placement test results and the students' scores of TOEFL ITP is checked. As it can be seen from Table 19 below, there is a significant relationship only between the average of reader results and TOEFL, $r = .44$ $n = 101$, $p > .05$, two-tailed test,

The correlation between these two variables and TOEFL scores for AF Level can be found below:

Table 19: *Correlations for AF Level Reader Exam, Placement test and TOEFL ITP*

		AFPLACE	AFRTOTAL	AFTOEFL
AFPLACE	Pearson Correlation		-.05	**.06**
	Sig. (2-tailed)		.65	**.56**
	N		101	**101**
AFRTOTAL	Pearson Correlation			**.44**[**]
	Sig. (2-tailed)			**.00**
	N			**101**

**. Correlation is significant at the 0.01 level (2-tailed).

As it can be seen in Table 20 below, there is a positive significant relationship between the variables and TOEFL ITP scores of students F $(24,76)$ = 7.38, p = 0.00.

Table 20*: ANOVA Table for AF Level*

Model		SS	Df	Mean Square	F	Sig.
1	Regression	64002.24	24	2666.76	7.38	.00[b]
	Residual	27482.95	76	361.62		
	Total	91485.19	100			

a. Dependent Variable: AFTOEFL
b. Predictors: (Constant). AFRTOTAL. AFPLACE. AFQVIII. AFQI. AFMTX. AFQXII. AFQIII. AFQIV. AFQX. AFQV. AFQVII. AFQXI. AFQII. AFMTVII. AFMTIII. AFQVI. AFQIX. AFMTI. AFMTIX. AFMTV. AFMTII. AFMTVIII. AFMTIV. AFMTVI

In the figures below, for the three highest correlations (between Quiz-9 (AFQIX) and TOEFL, Midterm-8 (AFMTVIII) and TOEFL, and overall reader average (AFRTOTAL) and TOEFL), scatter plots are drawn. As it can be seen in all three figures, there appears a positive strong linear relationship between these variables.
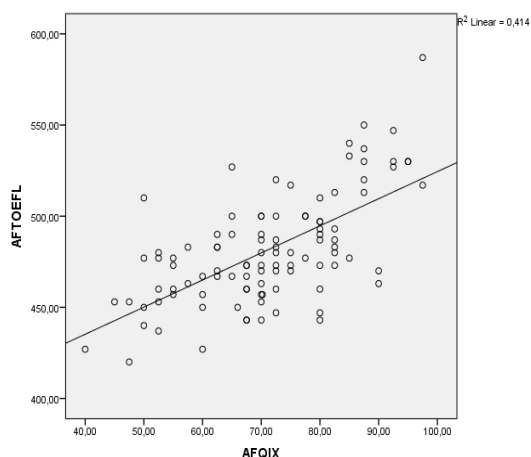
Figure 12: Scatter Plot of Students'
Quiz-9 Results and TOEFL
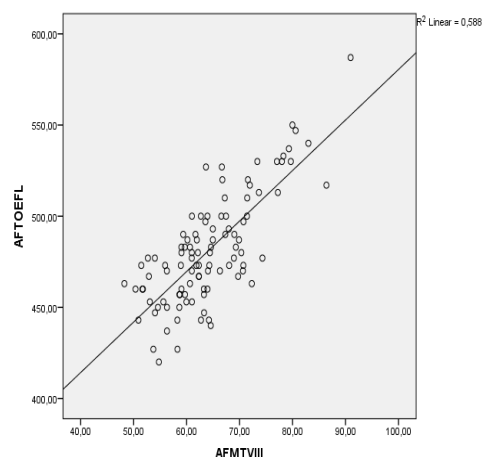
Figure 13: Scatter Plot of Students'
Midterm-8 results and TOEFL
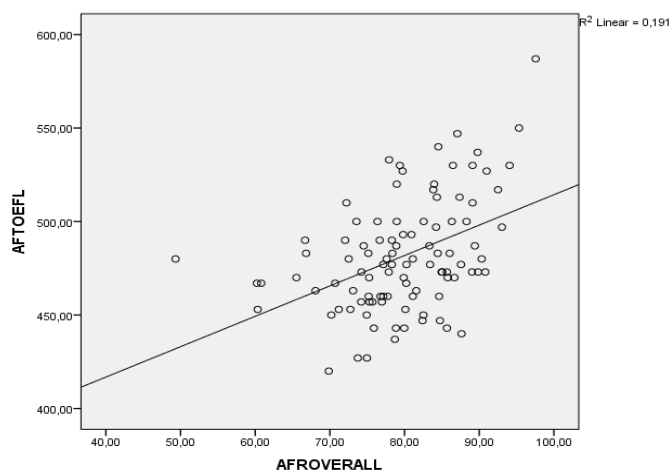
Figure 14: Scatter Plot of Students' Overall Reader Results and TOEFL

In addition to checking the relationship between the variables and TOEFL ITP scores one by one, the relationship between the combination of all the variables and TOEFL ITP is checked for AF Level. In Table 21 below, the magnitude of correlation coefficient (.70) indicates that students' TOEFL ITP scores are related to the all the variables to a high extent. More specifically, it can be said that these variables can explain 70% of the variation in TOEFL scores of the students.

Table 21: *Model Summary for AF Level*

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.84[a]	.70	.61	19.02

a. Predictors: (Constant). AFMTV. AFPLACE. AFQVIII. AFQI. AFQXII. AFMTX. AFRTOTAL. AFQIV. AFQIII. AFQX. AFQV. AFQVII. AFQXI. AFQII. AFMTVII. AFMTIII. AFQVI. AFQIX. AFMTI. AFMTIX. AFMTII. AFMTVIII. AFMTIV. AFMTVI

4.2.3. A Level (Elementary Level Students)

A Level refers to elementary level students whose English proficiency is higher than AF Level but lower than B Level. In Table 22 below, the descriptive analysis of the data in relation to all the assessment practices for A Level presents that among the quizzes while Quiz-1 (AQI) has the highest mean *(M= 83.82, SD= 8.16),* Quiz-10 (AQX) has the lowest mean *(M= 66.82, SD= 10.55).* Among the midterms, Midterm-8 (AMTVIII) has the lowest mean *(M= 67.69, SD= 8.67)* and Midterm-1 (AMTI) has the highest mean *(M= 78.70, SD= 7.35).* Moreover, while students' placement test grades range between 49.60 and 64.00, with an average of 56.68, their TOEFL scores range between 410 and 563, with an average of 490.70.

Table 22: *Descriptive Statistics for A Level*

	N	Minimum	Maximum	M	SD
APREL	**107**	**49.60**	**64.00**	**56.68**	**4.44**
AQI	**107**	**50.00**	**97.50**	**83.82**	**8.16**
AQII	107	69.25	96.63	83.12	6.53
AQIII	107	56.00	100.00	81.72	10.79
AQIV	107	50.00	95.00	75.72	9.62
AQV	107	60.00	92.50	75.49	8.07
AQVI	107	44.50	92.50	70.27	9.29
AQVII	107	53.50	100.00	79.07	11.84
AQVIII	107	42.50	95.00	72.29	11.69
AQIX	107	50.25	93.25	73.45	10.69
AQX	**107**	**35.00**	**92.50**	**66.82**	**10.55**
AQXI	107	55.00	100.00	79.80	11.26
AQXII	107	39.50	98.38	68.27	11.21
AMTI	**107**	**51.30**	**93.70**	**78.70**	**7.35**
AMTII	107	34.80	93.00	76.71	8.58
AMTIII	107	52.90	93.80	75.32	9.29
AMTIV	107	58.10	92.40	76.81	8.65
AMTV	107	57.10	94.40	75.79	8.27
AMTVI	107	47.00	90.20	72.26	10.46
AMTVII	107	55.50	93.60	73.87	9.22
AMTVIII	**107**	**49.67**	**89.69**	**67.69**	**8.67**
AMTIX	107	53.45	88.07	68.90	8.91
AMTX	107	18.48	90.63	69.13	9.89
ARTOTAL	107	61.17	97.63	80.44	8.04
ATOEFL	**107**	**410.00**	**563.00**	**490.69**	**31.26**
Valid N (listwise)	107				

To see the relationship between the variables one by one for A Level, Pearson Correlation was used and only the correlation between the variables and TOEFL Scores is analyzed. First of all, the relationship between all the quizzes and the students' scores of TOEFL ITP is checked for A Level. As it can be seen from Table 23 below, there is a significant relationship between all the quiz scores of the students and their TOEFL ITP results. While the highest significant correlation is

found between Quiz-10 and TOEFL, $r = .69$ $n = 107$, $p > .05$, two-tailed test, the lowest but again significant correlation is found between Quiz-3 results and TOEFL, $r = .26$, $n = 107$, $p > .05$, two-tailed test.

Secondly, the relationship between all the midterms and the students' scores of TOEFL ITP is checked for A Level. As it can be seen from Table 24 below, there is again a significant relationship between all the midterms and TOEFL. While the highest significant correlation is found between Midterm-9 and TOEFL, $r = .75$ $n = 107$, $p > .05$, two-tailed test, the lowest but again significant correlation is found between Midterm-1 and TOEFL, $r = .48$, $n = 107$, $p > .05$, two-tailed test.

In the table 23 and 24, correlation between quizzes and TOEFL scores for A Level and correlation between midterms and TOEFL scores for the same level can be found respectively.

Table 23: *Correlations for A Level Quizzes and TOEFL ITP*

		AQI	AQII	AQIII	AQIV	AQV	AQVI	AQVII	AQVIII	AQIX	AQX	AQXI	AQXII	ATOEFL
AQI	Pearson Correlation		.56**	.20*	.49**	.38**	.44**	.28**	.42**	.49**	.43**	.34**	.41**	**.48****
	Sig. (2-tailed)		.00	.04	.00	.00	.00	.00	.00	.00	.00	.00	.00	**.00**
	N		107	107	107	107	107	107	107	107	107	107	107	**107**
AQII	Pearson Correlation			.19*	.39**	.33**	.42**	.30**	.33**	.43**	.30**	.38**	.39**	**.36****
	Sig. (2-tailed)			.05	.00	.00	.00	.00	.00	.00	.00	.00	.00	**.00**
	N			107	107	107	107	107	107	107	107	107	107	**107**
AQIII	Pearson Correlation				.17	.15	.22*	.41**	.12	.07	.20*	.35**	.16	**.26****
	Sig. (2-tailed)				.09	.14	.02	.00	.21	.50	.04	.00	.11	**.01**
	N				107	107	107	107	107	107	107	107	107	**107**
AQIV	Pearson Correlation					.57**	.58**	.28**	.53**	.56**	.42**	.30**	.47**	**.55****
	Sig. (2-tailed)					.00	.00	.00	.00	.00	.00	.00	.00	**.00**
	N					107	107	107	107	107	107	107	107	**107**
AQV	Pearson Correlation						.50**	.23*	.39**	.42**	.42**	.33**	.46**	**.46****
	Sig. (2-tailed)						.00	.02	.00	.00	.00	.00	.00	**.00**
	N						107	107	107	107	107	107	107	**107**
AQVI	Pearson Correlation							.22*	.57**	.53**	.53**	.41**	.58**	**.68****
	Sig. (2-tailed)							.03	.00	.00	.00	.00	.00	**.00**
	N							107	107	107	107	107	107	**107**
AQVII	Pearson Correlation								.23*	.26**	.22*	.48**	.23*	**.31****
	Sig. (2-tailed)								.02	.01	.02	.00	.02	**.00**
	N								107	107	107	107	107	**107**
AQVIII	Pearson Correlation									.62**	.55**	.27**	.53**	**.66****
	Sig. (2-tailed)									.00	.00	.01	.00	**.00**
	N									107	107	107	107	**107**
AQIX	Pearson Correlation										.60**	.39**	.57**	**.64****
	Sig. (2-tailed)										.00	.00	.00	**.00**
	N										107	107	107	**107**
AQX	Pearson Correlation											.29**	.57**	**.69****
	Sig. (2-tailed)											.00	.00	**.00**
	N											107	107	**107**
AQXI	Pearson Correlation												.40**	**.38****
	Sig. (2-tailed)												.00	**.00**
	N												107	**107**
AQXII	Pearson Correlation													**.69****
	Sig. (2-tailed)													**.00**
	N													**107**

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

54

Table 24: *Correlations for A Level Midterms and TOEFL ITP*

		AMTI	AMTII	AMTIII	AMTIV	AMTV	AMTVI	AMTVII	AMTVIII	AMTIX	AMTX	ATOEFL
AMTI	Pearson Correlation	1	.560**	.579**	.661**	.575**	.566**	.517**	.639**	.581**	.545**	**.480****
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.000	.000	.000	**.000**
	N		107	107	107	107	107	107	107	107	107	**107**
AMTII	Pearson Correlation			.451**	.640**	.634**	.643**	.605**	.677**	.604**	.512**	**.587****
	Sig. (2-tailed)			.000	.000	.000	.000	.000	.000	.000	.000	**.000**
	N			107	107	107	107	107	107	107	107	**107**
AMTIII	Pearson Correlation				.557**	.597**	.541**	.564**	.521**	.547**	.507**	**.524****
	Sig. (2-tailed)				.000	.000	.000	.000	.000	.000	.000	**.000**
	N				107	107	107	107	107	107	107	**107**
AMTIV	Pearson Correlation					.746**	.689**	.628**	.726**	.681**	.619**	**.629****
	Sig. (2-tailed)					.000	.000	.000	.000	.000	.000	**.000**
	N					107	107	107	107	107	107	**107**
AMTV	Pearson Correlation						.786**	.696**	.721**	.691**	.587**	**.737****
	Sig. (2-tailed)						.000	.000	.000	.000	.000	**.000**
	N						107	107	107	107	107	**107**
AMTVI	Pearson Correlation							.723**	.709**	.729**	.637**	**.706****
	Sig. (2-tailed)							.000	.000	.000	.000	**.000**
	N							107	107	107	107	**107**
AMTVII	Pearson Correlation								.649**	.636**	.540**	**.606****
	Sig. (2-tailed)								.000	.000	.000	**.000**
	N								107	107	107	**107**
AMTVIII	Pearson Correlation									.806**	.675**	**.742****
	Sig. (2-tailed)									.000	.000	**.000**
	N									107	107	**107**
AMTIX	Pearson Correlation										.739**	**.747****
	Sig. (2-tailed)										.00	**.00**
	N										107	**107**
AMTX	Pearson Correlation											**.66****
	Sig. (2-tailed)											**.00**
	N											**107**

**. Correlation is significant at the 0.01 level (2-tailed).

As for the last step, the relationship between the average of readers, placement test results and the students' scores of TOEFL ITP is checked for A Level. As it can be seen from Table 25 below, there is a significant relationship only between the average of reader results and TOEFL, $r = .42$ $n = 107$, $p > .05$, two-tailed test.

The correlation between these three variables and TOEFL scores for A Level can be found below:

Table 25: *Correlations for A Level Reader Exam, Placement test and TOEFL ITP*

		APLACE	ARTOTAL	ATOEFL
	Pearson Correlation		.12	**.16**
APLACE	Sig. (2-tailed)		.21	**.10**
	N		107	**107**
	Pearson Correlation			**.42**[**]
ARTOTAL	Sig. (2-tailed)			**.00**
	N			**107**

**. Correlation is significant at the 0.01 level (2-tailed).

In the figures below, for the three highest correlations for each part, scatter plots are drawn. Therefore, in the figures below, the correlation between Quiz-10 (AQX) and TOEFL, Midterm-9 (AMTIX) and TOEFL, and total reader results (ARTOTAL) and TOEFL can be seen. As it can be seen in all three figures, there appears a positive strong linear relationship between these variables.
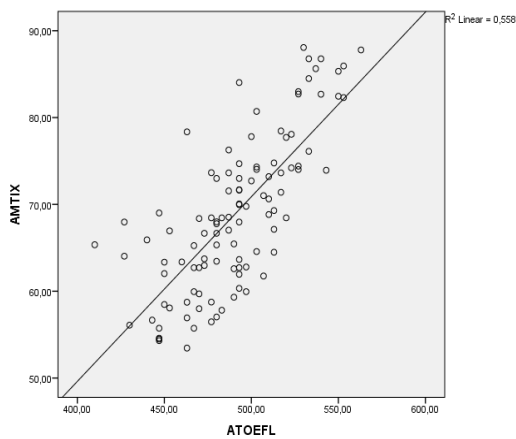
Figure 15: Scatter Plot of Students'
Midterm-9 Results and TOEFL
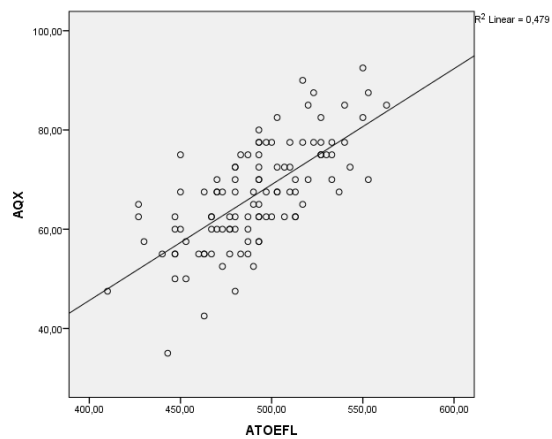
Figure 16: Scatter Plot of Students'
Quiz-10 results and TOEFL
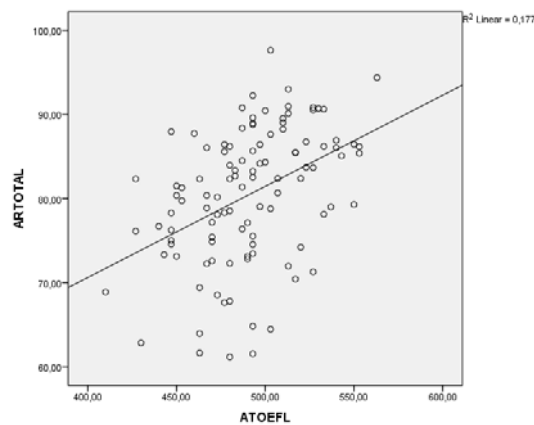
Figure 17: Scatter Plot of Students' Overall Reader Results and TOEFL ITP

In addition to checking the relationship between the variables and TOEFL ITP scores one by one, the relationship between the combination of all the variables and TOEFL ITP is checked for A Level. In Table 26 below, the magnitude of correlation coefficient (.78) indicates that students' TOEFL ITP scores are related to the all the variables to a high extent. More specifically, it can be said that these variables can explain more than half of the variation, exactly 70% in TOEFL scores.

Table 26: *Model summary for A Level*

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.88[a]	.78	.71	16.87

a. Predictors: (Constant), ARTOTAL, AQIII, APLACE, AQVIII, AQVII, AQV, AQII, AQXI, AQX, AMTIII, AMTX, AQI, AQVI, AQIV, AMTVII, AQIX, AQXII, AMTI, AMTII, AMTVIII, AMTIV, AMTVI, AMTV, AMTIX

As it can be seen in Table 27 below, there is a positive significant relationship between the variables and TOEFL ITP scores of students $F_{(24,82)} = 11.75$, p = 0.00.

Table 27: *ANOVA Table for A Level*

Model		SS	Df	Mean Square	F	Sig.
	Regression	80253.11	24	3343.88	11.75	.00[b]
1	Residual	23341.71	82	284.66		
	Total	103594.82	106			

a. Dependent Variable: ATOEFL

b. Predictors: (Constant). ARTOTAL. AQIII. APLACE. AQVIII. AQVII. AQV. AQII. AQXI. AQX. AMTI. AQI. AQXII. AQVI. AMTII. AMTVII. AMX. AQIX. AQIV. AMTIII. AMTVIII. AMTVI. AMTIV. AMTV. AMTIX

4.2.4. B Level (Intermediate Level Students)

B Level refers to intermediate level students who have higher English proficiency than both AF and A Level. However, like A and AF Level students, they are also expected to take the TOEFL test at the beginning of July if they meet the requirements mentioned before. The descriptive analysis of the data in relation to all the assessment practices for B Level presents that among the quizzes while Quiz-10 (BQX) has the highest mean *(M= 86.04, SD= 10.16),* Quiz-6 (BQVI) has the lowest mean *(M= 66.54, SD= 9.20).* Among the midterms, Midterm-1 (BMTI) has the lowest mean *(M= 69.49, SD= 7.53)* and Midterm-3 (BMTIII) has the highest mean *(M= 79.54, SD= 8.18).* While students' placement test grades range between 64.80 and 94.40, with an average of 76.59, their TOEFL scores range between 463 and 623, with an average of 524.95.

Table 28: *Descriptive Statistics for B Level*

	N	Minimum	Maximum	M	SD
BPLACE	**129**	**64.80**	**94.40**	**76.59**	**6.76574**
BQI	129	47.50	95.00	72.56	10.06
BQII	129	47.75	89.88	71.43	9.04
BQIII	129	56.00	100.00	81.54	10.51
BQIV	129	40.00	97.50	69.46	11.70
BQV	129	55.00	95.00	78.35	8.57
BQVI	**129**	**40.00**	**91.00**	**66.54**	**9.20**
BQVII	129	51.50	100.00	81.85	11.36
BQVIII	129	44.00	95.00	71.97	10.25
BQIX	129	50.38	99.25	77.58	11.30
BQX	**129**	**53.00**	**100.00**	**86.04**	**10.16**
BQXI	129	39.50	95.00	67.87	12.14
BQXII	129	42.63	98.75	70.84	11.62
BMTI	**129**	**47.90**	**88.40**	**69.49**	**7.53**
BMTII	129	51.70	92.40	74.58	7.67
BMTIII	**129**	**56.30**	**96.00**	**79.54**	**8.18**
BMTIV	129	53.50	91.80	72.60	8.38
BMTV	129	51.70	95.40	77.56	8.66
BMTVI	129	52.90	94.10	72.30	8.66
BMTVII	129	52.90	93.10	75.08	8.78
BMTVIII	129	60.53	95.66	78.46	7.62
BMTIX	129	58.91	95.05	78.52	7.34
BMTX	129	57.23	97.03	78.44	8.07
BRTOTAL	129	63.79	96.75	84.56	7.95
BTOEFL	**129**	**463.00**	**623.00**	**524.95**	**28.33**
Valid N (listwise)	129				

To see the relationship between the variables one by one for B Level, Pearson Correlation was used, and only the correlation between the variables and TOEFL Scores is analyzed. First of all, the relationship between all the quizzes and the students' scores of TOEFL ITP is checked for B Level. As it can be seen from Table 29 below, there is a significant relationship between all the quiz scores of the students and their TOEFL ITP results. While the highest significant correlation is

found between Quiz-12 (BQXII) and TOEFL, $r = .63$ $n = 129$, $p > .05$, two-tailed test, the lowest but again significant correlation is found between Quiz-3 (BQIII) results and TOEFL, $r = 24$, $n = 129$, $p > .05$, two-tailed test.

Secondly, the relationship between all the midterms and the students' scores of TOEFL ITP is checked for B Level. As it can be seen from Table 30 below, there is a significant relationship between all the midterms and TOEFL. While the highest significant correlation is found between Midterm-9 (BMIX) and TOEFL, $r = .74$ $n = 129$, $p > .05$, two-tailed test, the lowest but again significant correlation is found between Midterm-1 (BMI) and TOEFL, $r = .51$, $n = 129$, $p > .05$, two-tailed test.

In Tables 29 and 30 below, correlation between midterms and TOEFL scores for B Level and correlation between quizzes and TOEFL scores for the same level can be found respectively.

Table 29: *Correlations for B Level Quizzes and TOEFL ITP*

		BQI	BQII	BQIII	BQIV	BQV	BQVI	BQVII	BQVIII	BQIX	BQX	BQXI	BQXII	TOEFL
BQI	Pearson Correlation		,419**	,350**	,368**	,297**	,205*	,309**	,285**	,251**	,296**	,282**	,296**	**,366****
	Sig. (2-tailed)		,000	,000	,000	,001	,020	,000	,001	,004	,001	,001	,001	**,000**
	N		129	129	129	129	129	129	129	129	129	129	129	**129**
BQII	Pearson Correlation			,325**	,586**	,457**	,469**	,268**	,412**	,487**	,175*	,519**	,408**	**,395****
	Sig. (2-tailed)			,000	,000	,000	,000	,002	,000	,000	,047	,000	,000	**,000**
	N			129	129	129	129	129	129	129	129	129	129	**129**
BQIII	Pearson Correlation				,317**	,136	,136	,536**	,200*	,117	,547**	,203*	,102	**,243****
	Sig. (2-tailed)				,000	,124	,124	,000	,023	,185	,000	,021	,248	**,005**
	N				129	129	129	129	129	129	129	129	129	**129**
BQIV	Pearson Correlation					,510**	,518**	,279**	,559**	,517**	,318**	,591**	,489**	**,472****
	Sig. (2-tailed)					,000	,000	,001	,000	,000	,000	,000	,000	**,000**
	N					129	129	129	129	129	129	129	129	**129**
BQV	Pearson Correlation						,558**	,243**	,567**	,584**	,266**	,574**	,601**	**,552****
	Sig. (2-tailed)						,000	,005	,000	,000	,002	,000	,000	**,000**
	N						129	129	129	129	129	129	129	**129**
BQVI	Pearson Correlation							,195*	,554**	,570**	,113	,498**	,537**	**,595****
	Sig. (2-tailed)							,027	,000	,000	,203	,000	,000	**,000**
	N							129	129	129	129	129	129	**129**
BQVII	Pearson Correlation								,210*	,204*	,498**	,197*	,170	**,255****
	Sig. (2-tailed)								,017	,020	,000	,025	,055	**,004**
	N								129	129	129	129	129	**129**
BQVIII	Pearson Correlation									,606**	,220*	,594**	,568**	**,625****
	Sig. (2-tailed)									,000	,012	,000	,000	**,000**
	N									129	129	129	129	**129**
BQIX	Pearson Correlation										,272**	,623**	,615**	**,604****
	Sig. (2-tailed)										,002	,000	,000	**,000**
	N										129	129	129	**129**
BQX	Pearson Correlation											,272**	,132	**,273****
	Sig. (2-tailed)											,002	,136	**,002**
	N											129	129	**129**
BQXI	Pearson Correlation												,639**	**,546****
	Sig. (2-tailed)												,000	**,000**
	N												129	**129**
BQXII	Pearson Correlation													**,626****
	Sig. (2-tailed)													**,000**
	N													**129**

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Table 30: *Correlations for B Level Midterms and TOEFL ITP*

		BMTI	BMTII	BMTIII	BMTIV	BMTV	BMTVI	BMTVII	BMTVIII	BMTIX	BMTX	TOEFL
BMTI	Pearson Correlation	1	,62**	,64**	,59**	,57**	,50**	,52**	,54**	,49**	,44**	**,51****
	Sig. (2-tailed)		,00	,00	,00	,00	,00	,00	,00	,00	,00	**,00**
	N			129	129	129	129	129	129	129	129	**129**
BMTII	Pearson Correlation		1	,72**	,70**	,67**	,67**	,61**	,69**	,67**	,60**	**,55****
	Sig. (2-tailed)			,00	,00	,00	,00	,00	,00	,00	,00	**,00**
	N			129	129	129	129	129	129	129	129	**129**
BMTIII	Pearson Correlation			1	,64**	,72**	,69**	,64**	,62**	,63**	,54**	**,61****
	Sig. (2-tailed)				,00	,00	,00	,00	,00	,00	,00	**,00**
	N				129	129	129	129	129	129	129	**129**
BMTIV	Pearson Correlation				1	,74**	,72**	,74**	,75**	,71**	,65**	**,70****
	Sig. (2-tailed)					,00	,00	,00	,00	,00	,00	**,00**
	N					129	129	129	129	129	129	**129**
BMTV	Pearson Correlation					1	,74**	,71**	,71**	,70**	,66**	**,66****
	Sig. (2-tailed)						,00	,00	,00	,00	,00	**,00**
	N						129	129	129	129	129	**129**
BMTVI	Pearson Correlation						1	,71**	,74**	,74**	,71**	**,71****
	Sig. (2-tailed)							,00	,00	,00	,00	**,00**
	N							129	129	129	129	**129**
BMTVII	Pearson Correlation							1	,74**	,72**	,69**	**,65****
	Sig. (2-tailed)								,00	,00	,00	**,00**
	N								129	129	129	**129**
BMTVIII	Pearson Correlation								1	,80**	,75**	**,72****
	Sig. (2-tailed)									,00	,00	**,00**
	N									129	129	**129**
BMTIX	Pearson Correlation									1	,79**	**,74****
	Sig. (2-tailed)										,00	**,00**
	N										129	**129**
BMTX	Pearson Correlation										1	**,70****
	Sig. (2-tailed)											**,00**
	N											**129**

**. Correlation is significant at the 0.01 level (2-tailed).

As for the last step, the relationship between the average of readers, placement test results and the students' scores of TOEFL ITP is checked for B Level. As it can be seen from Table 31 below, there is a significant relationship between the average of reader results and TOEFL, $r = .33$ $n = 129$, $p > .05$, two-tailed test and between the students' placement test results and their TOEFL ITP scores, $r = .36$ $n = 129$, $p > .05$, two-tailed test.

The correlation between these three variables and TOEFL scores for B Level can be found below:

Table 31: *Correlations for B Level Reader Exam, Placement Test and TOEFL ITP*

		BPLACE	BRTOTAL	BTOEFL
BPLACE	Pearson Correlation		-.04	.36**
	Sig. (2-tailed)		.62	.00
	N		129	129
BRTOTAL	Pearson Correlation			.33**
	Sig. (2-tailed)			.000
	N			129

**. Correlation is significant at the 0.01 level (2-tailed).

In the figures below, for the three highest correlations for each part, scatter plots are drawn. The correlation between Quiz-12 (BQXII) and TOEFL, Midterm-9 (BMTIX) and TOEFL, and placement exam results and TOEFL can be seen. According to these three figures, there appears a positive strong linear relationship between these variables.
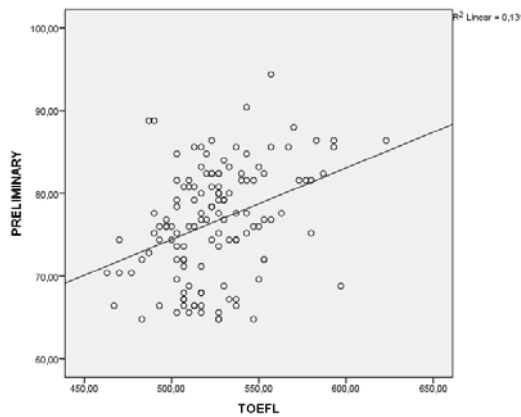
Figure 18: Scatter Plot of B Level Students'   *Figure 19:* Scatter Plot of

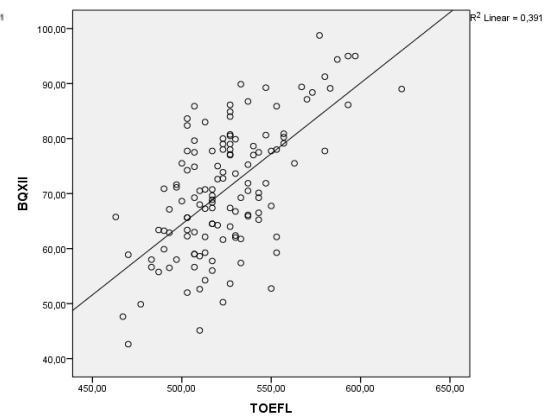Placement Exam Results and TOEFL   Quiz-12 Results and TOEFL
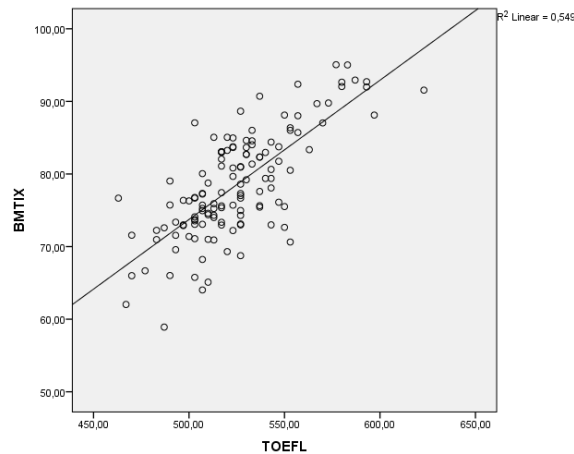
Figure 20: Scatter Plot of B Level Students' Midterm-9 Results and TOEFL

In addition to checking the relationship between the variables and TOEFL ITP scores one by one, the relationship between the combination of all the variables and TOEFL ITP is checked for B Level. In Table 35 below, the magnitude of correlation coefficient (.70) indicates that students' TOEFL ITP scores are related to the all the variables to a high extent. More specifically, it can be said that these variables can explain more than half of the variation, exactly 70%, in TOEFL scores.

Table 32: *Model summary for B Level*

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.84[a]	.70	.63	17.28

a. Predictors: (Constant), BRTOTAL, BQI, BQX, BQVI, BQV, BQVII, BQIII, BQII, BQVIII, BMTX, BQIV, BPLACE, BQXII, BMTIII, BQIX, BQXI, BMTI, BMTII, BMTVII, BMTIV, BMTV, BMTVI, BMTIX, BMTVIII

As it can be seen in Table 33 below, there is a positive significant relationship between the variables and TOEFL ITP scores of students $F_{(24,104)} = 10.18$, p = 0.00.

Table 33: *ANOVA Table for B Level*

Model		SS	Df	Mean Square	F	Sig.
	Regression	72070.93	24	3002.96	10.18	.000[b]
1	Residual	30692.79	104	295.12		
	Total	102763.72	128			

a. Dependent Variable: TOEFL

b. Predictors: (Constant). BRTOTAL. BQI. BQX. BQVI. BQVII. BQII. BQVIII. BQIII. BQV. BQIV. BMTX. BPLACE. BQXII. BQIX. BQXI. BMTII. BMTI. BMTIII. BMTVII. BMTIV. BMTV. BMTVI. BMTIX. BMTVIII

4.2 Results of the Qualitative Data

In this section of the study, the results of questionnaires given to instructors and students are discussed according to the variables of the study. For each category, examples related to the variable are provided. As a last step, a comparison between the results of student and instructor questionnaire is made.

4.2.1 Results of the Instructor Questionnaire

The instructors were given a questionnaire which had three open-ended questions. (See Appendix B) Each of these questions was related to a variable namely quizzes, midterms, and readers. In addition, they were asked to give a point out of six to show

the degree of the relationship between these assessment practices and TOEFL ITP. To begin with, in the table below, the average of points given to each assessment practice is provided.

Table 34: *The Results of the Instructor Questionnaire: Question 8, 10, and 12*

	Quizzes	*Midterms*	*Readers*
Mean (out of 6)	4.27	4.97	2.97

As it can be understood from the table above, the instructors think that while the highest correlation is expected to be found between midterms and TOEFL ITP scores of students, the lowest one is expected regarding the one between readers and TOEFL ITP.

4.2.1.1 Quizzes

As for the first question, the instructors were expected to share their opinion about the effect of quizzes on the success of TOEFL ITP. There were some common points in their responses, but there were some other comments which did not fall into any categories.

Out of 48 instructors, three did not make any comments related to quizzes. Out of 45 who shared their opinion with the help of the questionnaire, 32 are of the opinion that as quizzes, in a way, force students to study regularly, they help them to learn and get prepared better for the upcoming TOEFL test. As one of the instructor claimed:

I think there is definitely a positive correlation between quizzes and TOEFL ITP. As students are having a quiz almost every week, it motivates them to study regularly. After some time, they start to have less stress and regard such exams as a regular event. (Instructor6)

They are also thinking that unless the students are required to take quizzes on regular basis, they will not spend any time studying English. Therefore, some of them focused on the fact that by the help of quizzes, they have no option but to revise,

remember and learn the items which will be also covered in the upcoming test, TOEFL ITP. One of the instructors focused on the positive effect of having quizzes in relation with TOEFL by saying that:

First of all, before each quiz, students need to revise newly-learned topics and solve related questions. Right after the quiz, they are getting immediate response and being aware of their weaknesses. Therefore, having such a hectic schedule helps them to be alert all the time and finally to be more successful for the TOEFL ITP at the end of the semester. (Instructor34)

Eleven instructors compared the types of question asked in the quizzes and the ones in TOEFL ITP. Therefore, if a quiz consists of different types of questions rather than the one asked in TOEFL, they tend to regard it as less effective and powerful in predicting the success of students in TOEFL. There are some other comments which focus on the importance of having similar question types along with some suggestions. As two instructors said below:

When different types of questions are asked in the quizzes such as word formation or rewrite, I do not think they match with the format of TOEFL ITP, which makes it less powerful when compared with midterms. Therefore, it is very important to have quizzes including question types just like the ones asked in TOEFL ITP. By doing so, we can make our students get to know such types and help them obtain higher grades. (Instructor14)

Listening part and reading part in quizzes are exactly the same with that of TOEFL. However, for structure part, a productive part, where the students are expected to form or compose some sentences or key words on their own, is sometimes included. As the students are not expected to do so in TOEFL, I do not think such parts are necessary. (Instructor35)

For them, having same or parallel question types in quizzes are very useful in helping students have an idea about the question types asked in TOEFL ITP. Therefore, six of them think that when TOEFL ITP types of questions are included, it may look like a smaller version and help students to be ready for it. However, it is not possible to have only those types of questions in quizzes. Therefore, they believe that the effect of quizzes will be less than the midterms. Therefore, they suggest that especially in the third semester, same or similar types of questions need to be asked.

In addition to question types and regular study habits, many of the instructors also share their opinions on the content of the quizzes especially writing and speaking ones. These two skills are not covered in TOEFL ITP; therefore, five of the instructors claim that even if they are indispensable components of English, there is no significantly important relationship between these two. As explained in the comment:

The content of the quizzes is just limited to the newly-learned grammar topics in the last couple of weeks. In other words, quizzes just measure one or two specific grammar topics that have been covered lately, and therefore, the level of difficulty is too different to be compared with TOEFL test, I personally believe that because of this limitation, their effect on the success of students taking TOEFL test will be less than that of midterms. (Instructor7)

Out of 45, 13 instructors also suggested that both the number and the length of quizzes should be reduced. They believe that a stronger relationship can be found between them and TOEFL ITP only after such a change. As for the number, they think that having many quizzes are just making the students more stressful. Likewise, as for the length they believe that for students who are getting ready to take TOEFL ITP, a test which requires the students to be quick, giving them so much time for a quiz is something that needs to be changed. One of the instructors claimed that:

Quizzes help students to get used to taking exams; however, as the number is too many, students are just learning for the sake of getting good grades or in some other cases, it just creates exam anxiety. Therefore, in order to reduce it, we should definitely reduce the number of exams even though I personally believe they are very effective in getting students prepared for the upcoming test. (Instructor11)

Among all the comments, there is only one which was very different from the others. While all the other comments are relatively positive with one or two main issues or problems raised, the one, made by one of the instructors working at the institution less than a year, claims that the quizzes are totally ineffective and unnecessary.

The quizzes applied almost every week are doing nothing for our students except making them stressful and diminishing their motivation. (Instructor 44)

4.2.1.2 Midterms

As for the second question, the instructors were expected to share their opinion about the effect of midterms on the success of students for TOEFL ITP. Even though similar comments were made for this question, all of the instructors believe that a higher correlation will be found between the midterm scores and TOEFL ITP rather than the one between quizzes and TOEFL test. Moreover, as the last three midterms are just copies of official TOEFL ITP both in terms of the number of questions and the content, 34 instructors out 45 who made comments related to the midterms had the opinion that the last three had a higher positive correlation to TOEFL. The teachers' responses can also be categorized under different headings. To begin with, the teachers compare the quizzes to midterms, and talked about their relationship between TOEFL. For various reasons, one of which is wide range of topics that are covered in this assessment practice, they thought that predictive power of midterms will be greater than that of quizzes. While the quizzes cover only the newly-learnt topics in the last two weeks, midterms cover a wide range of topics. Therefore, 29 of them are of the opinion that the more subjects are covered, the more useful they will become for getting higher scores from TOEFL ITP. As one claims:

Midterms are more comprehensive than quizzes in terms of both the variety of questions and the topics covered. Even though it might be long and tiring for students, having numerous questions related to different topics is very effective for the success of students. (Instructor5)

Ten teachers compared the limited number of skills covered in quizzes with the ones included in midterms. For them, as in the quizzes, listening and writing are asked interchangeably, which means just one at a time, in midterms all the skills and areas are covered. Therefore, they believe that it is a better practice for upcoming TOEFL test.

Moreover, there are some key points which are frequently asked in TOEFL. Especially in listening parts, idioms are one of the items that the students should be

familiar with. Therefore, that these idioms are asked in the assessment practices especially in midterms is found very effective by the instructors. As one claims:

When the content of the midterms are similar to that of TOEFL, I believe that the students are getting the most benefit out of it. For example, the students are trying to learn idioms as they are asked in the listening part of the midterms. However, as they are also asked in the official TOEFL ITP, it becomes very useful for the students. (Instructor27)

While in quizzes, there might be different types rather than the ones asked in TOEFL such as rewrite or fill in the blanks with one word, in midterms the question types are similar to the ones in TOEFL. Therefore, thirty four teachers agree that by having similar questions namely multiple choice and error recognition, types used in the TOEFL test, we are making midterms more powerful in terms of making students getting ready for TOEFL. Therefore, they believe that midterms help our students to have practice and get ready better for it.

As mentioned before, almost all of the teachers separate the last three midterms from the others. These midterms are exactly like TOEFL in terms of the duration, content and the types of questions. Therefore, they are expecting to find a higher correlation between them and the official TOEFL ITP. Some of them also suggested that having more TOEFL-like midterms might be very beneficial. One of the instructor suggested that by saying:

I do not think the midterms in the first term are very effective in getting higher scores from TOEFL but rather learning general English. However, in the last term, especially the last three midterms were very effective as they are just like TOEFL test. (Instructor48)

Another issue which is raised by the instructors is timing and/or duration of the quizzes. While the quizzes approximately take about 50 minutes, midterms last more than two hours. As the official TOEFL test also takes a little bit more than two hours, five of the instructors are of the opinion that having similar duration to TOEFL test makes midterms more effective and powerful in predicting students' scores:

As midterms are longer than quizzes, it becomes a very good practice for our students who may have some concentration problems. However, they have to overcome this problem as in the upcoming test, they are also expected to work for about two hours. That's why I personally believe that having similar time limit makes midterm more effective than quizzes. (Instructor31)

The other similarity between the midterms and the TOEFL test is related to specific time limit for each part of the test. For example, as the time given to the structure part is 25 minutes, the students are only allowed to work on this section during that time. When the time is up, they can work on the next section without going back to the previous section. This rule is also applied to our midterms, which makes them more similar to TOEFL and therefore, more effective according to some instructors. According to one of the instructor:

Having separate time limit for each part of the exam, namely listening, structure and reading, is very effective as most of our students were not aware of such a fact. Even though they had some problems obeying this rule at the outset, they became more competent as the time passed and this made their job much easier in the official TOEFL test. (Instructor11)

Four teachers also shared their opinion about the positive effect of midterms in general. For them, as the students are getting used to the exams, they are having less stress, which makes them more successful. Likewise, they think that with the help of midterms the students are aware of their level and progress. Getting feedback after each midterm is not only beneficial for the students but also for the teachers who may work hard to help his students revise the problematic parts:

Even though for the quizzes, some of the students may just try to memorize and get higher grades, for midterms, it might not be possible to do same thing as they cover more. Therefore, one of the positive effect of having midterms are the fact that it hinders memorization and forces students to revise, comprehend and finally internalize the subjects. (Instructor7)

4.2.1.3 Reader Exams

The last type of exam carried out at TOBB ETU is reader exam, which requires students to read one or two pre-determined book for each reader exam and answer

related questions. Even though the teachers think that the idea of having reader exams is very beneficial, our students are not getting the most out it. Therefore, out of 46 who had comments related to reader exams, 30 of them have the opinion that reader exams are not useful or effective.

For each reader exam, not only comprehension questions but also some vocabulary questions are asked. These questions may involve matching the key words with their definitions, putting the word in correct places, finding synonyms or antonyms of the key words which appeared in the pre-determined books. In lower levels, students may be asked to match the key words with the pictures. Therefore, four of the instructors believe that it is useful for vocabulary development.

Three instructors find readers quite useful as most of them are national geographic reader series. Therefore, the topic may be the same as the one that they will encounter in the official TOEFL ITP. On the other hand, even though the instructors do not believe that there is a direct correlation between the readers, two of them claim that some of their students get very curious about the topics of the books and as a result, they read them in detail:

As each reader is about a different topic which might interest some of our students, I believe that they get familiar with different areas. Consequently, in the official TOEFL ITP, they may encounter a similar one and read it with more confidence. (Instructor1)

In addition the problems related to those exams, twelve instructors believe that even though there might not be a direct correlation between the reader exams and TOEFL, they are helping our students to develop their reading skills. Especially, if the students read the pre-determined books by keeping the TOEFL strategies in mind, the instructors believe that it would be more beneficial:

If a student reads the book in detail, revises the vocabulary items, writes a summary, I highly believe that it will be very useful for the development of his reading skill. (Instructor9)

Twenty seven of the teachers believe that more than half of our students are not reading the book but just getting the summary from their friends or making them give some details so that they can learn the flow of friends and get a passing grade:

In the long term, it might be effective; however, our students are just reading for the sake of exams. For them, knowing the name of the characters and the flow of events along with the key words are more than enough. Therefore, rather than spending time on it, asking a friend to summarize it for them in Turkish, of course, just before the exam will be fine. (Instructor8)

All they do is to find a friend who has read the book - the number of students reading it is very low- and ask him/her to summarize the important point. Therefore, just on the day of the time, maybe just one hour before it, making him talk about specific points will be enough to get a passing grade. (Instructor333)

Even though they do believe that reading is one of the most important skills and we have to do our best to make our students read, they do not think that having a reader exam is the best way of doing it. Even though they believe that reader exams may encourage their students to read outside the classroom and become a fast reader, they do not think that they fulfill their aims. The students are required to take an exam afterwards, which just forces them to memorize and get a higher grade. Therefore, they did not find them effective and five of them suggested other ways that might be useful for our students. For them, it is not the exams which help them to develop their reading skills but the books which they are required to read. Therefore, they demanded that rather than having an exam afterwards, the students be asked to carry out different tasks in class. Some suggestions are given by two instructors:

Instead of having reader exams which do not fulfill their purpose, it might be a better idea to ask our students to prepare a portfolio task related to the pre-determined book. (Instructor20)
I think a different assessment practice might be more useful as they are not reading the whole book. Instead, they find a friend to summarize it for them. Therefore, they may prepare a game and present it to the whole class related to the book that they are supposed to read. (Instructor2)

4.2.2 Results of the Student Questionnaire

Like instructors, students were also given a questionnaire which consists of three open-ended questions, each of which is related to a variable namely quizzes, midterms, and readers (See Appendix A). In addition, they are asked to give a point out of six to show the degree of the relationship between these assessment practices and TOEFL ITP. However, out of 421, 34 of them did not make any comments only gave a point to indicate the degree of relationship between the assessment practices and TOEFL ITP. Therefore, the answers given by 387 students were analyzed. To begin with, in the table below, the average of points given to each assessment practice is provided.

Table 35: *The Results of the Student Questionnaire: Questions 16, 18, and 20*

	Quizzes	Midterms	Readers
Mean (out of 6)	4.01	4.98	2.89

As it can be understood from the table above, students think that while the highest correlation is expected to be found between midterms and TOEFL ITP scores of students, the lowest one is expected regarding the one between readers and TOEFL ITP.

4.3.2.1 Quizzes

As for the first question, the students were expected to share their opinion about the effect of quizzes for the success of TOEFL ITP. Their responses can be categorized under different headings. For this part, this categorization is done with the help of two main titles.

Positive comments

Out of 387, 221 students have a positive attitude towards the quizzes. They believe that thanks to quizzes, they feel obliged to revise the subjects that are necessary for getting higher scores from TOEFL. However, they also think that there is not a direct

relationship between these two but rather quizzes being necessary for preparing the path towards the TOEFL test. In other words, whether the students have completely covered the newly-learnt topics or not is checked by the help of quizzes:

Each and every quiz is useful for our improvement as we have feedback after each quiz and understand the parts that we haven't got right. If we hadn't had that many quizzes, I would not have studied in a regular basis, therefore, quizzes, in a way, force us to study and revise all the time. (Student98)

The students believe that with the help of the some parts of quizzes, they had a chance to get used to solving questions especially in reading, listening and structure sections. Therefore, they consider these exams as a chance to practice some of the important skills especially listening and structure. For them, quizzes are making them less stressful as they are getting used to them day by day along with helping them to get better at some of the question types and necessary skills and learning their weaknesses. According to one student:

By the help of quizzes, I know which part I am very good at and which parts I need to improve. They are helping us to be more successful students throughout the term and at the end they are ready to get higher grades from the TOEFL test. (Student387)

Out of 221 students who have a positive attitude towards quizzes, 176 of them believe that quizzes are very useful for practicing specific skills especially listening and reading. As they think the types of questions are like the ones asked in TOEFL such as short conversations related to idioms, long conversation or lectures, they believe that they are very effective for getting higher scores from these parts.

I think the most effective part of quizzes is the listening. The questions are just like the ones asked in TOEFL test. If a student can answer the questions in the listening sections of our tests, I believe he can also be successful in the listening section of the official test. (Student291)

Negative comments

Out of 387, 166 of them have made negative comments related to the predictive power of quizzes to TOEFL ITP. They think that because of some question types

which are unrelated to official TOEFL ITP, in other words which are not asked in that test, quizzes are not effective for getting higher points from the TOEFL test. Therefore, they believe that if the format of some of the quizzes is changed and similar questions to TOEFL such as multiple choice and error recognition are asked, they believe that quizzes will be more effective and useful:

Sometimes, they ask some re-write or fill-in-blanks types of questions in the quizzes. However, in TOEFL, we are not going to have such parts. Therefore, we should have more multiple choice and error recognition types of questions. (Student4)

The students also believe that as there is no speaking, writing and vocabulary sections in the official TOEFL ITP, such parts are unnecessary and ineffective. Therefore, they claim that as for the vocabulary part of the quizzes, they just try to memorize the key words that are highly likely to be asked. Likewise, they do not think that they need to have writing or speaking part in these exams. For them, these assessment practices are just making them feel stressed:

We have a speaking quiz and some of our quizzes have a writing section, but these skills are not assessed in the official test. Therefore, like most of my friends, I believe that they are not necessary. (Student315)

For some students, the level of quizzes is not appropriate for them as they are getting ready for an exam which is internationally recognized. As one of the students claimed:

I think that the quizzes are below our level. That is why they are less effective than midterms. (Student4)

The next issue raised by the students is the fact that in the quizzes limited content is covered. Therefore, some claimed that just for the sake of getting good grades, they are revising the subjects that are covered in the last couple of weeks. As there is only a limited number of subjects that need to be revised and learned, they think that quizzes are less effective than midterms which cover more and require more effort:

We are only asked questions related to the topic of previous week, but in TOEFL there are so many subjects that we need to know. Therefore, I do not think they are as effective as midterms. (Student113)

4.2.2.2 Midterms

Like instructors, students also believe that the highest correlation is to be found between midterms and TOEFL because of the similarity of such tests to TOEFL. The students categorize these similarities as the hardness, duration and types of questions. They think that just in one exam, it is possible to have an idea of all the skills that the TOEFL test requires; therefore, of out 387, 298 students are of opinion that midterms are more effective than quizzes for our success at this test.

TOEFL is a long test which takes about two hours. At first, I thought I could not focus for such a long time. However, by the help of midterms, we get accustomed to staying focused for two or more hours. (Student39)

The students also shared their opinions on the last three midterms. They believe that these different assessment practices are not so good except the ones which are true copies of official TOEFL ITP. Therefore, some students ($n = 56$) suggest that midterms with unrelated parts to TOEFL should be eliminated and different ones with similar question types and parts should be formed. Likewise, as the number of questions in midterms except last three is not the same as TOEFL, they think there is a need for a change. In short, for them having three exact TOEFL-type midterms is a great practice and very effective in their success for TOEFL test. As two students claimed:

I personally believe that the most effective assessment practice is the last three midterms. They have the same number of questions, the same format and time given is also the same. We should have more of that kind. Apart from that, I do believe that the other midterms are not so good. (Student53)

Why do we have a writing section when there is not writing part in official TOEFL ITP or why are they asking productive type of questions. I think these unnecessary parts should be omitted and error recognition and multiple type of questions should be asked instead. (Student67)

4.2.2.3 Reader Exams

The students also made different comments related to reader exams. While out of 387 students who made comments related to these exams, 124 find them quite useful, 263 students do not think they are as effective as they can be. The first issue related to readers is the fact that as the topics of these pre-determined books might be related or similar to the topics of the reading passages in TOEFL exam, they find them effective. They believe that by the help of readers, they are learning some new vocabulary items which might be useful for them in the upcoming test:

I think they are effective for the reading passages asked in TOEFL test. There might be a chance that readings with similar topics can be asked. (Student18)

I learned different key words by the help of these books and I had not encountered such words even within the classroom environment. (Student47)

In the TOEFL test, the students are required to read five passages and answer 50 questions within 55 minutes. Therefore, being a fast reader is very important for getting higher scores. They believe that along with improving reading skills and vocabulary knowledge, reading these books help them to become fast readers:

Even though some of the books chosen for the reader exam is very long, I try to read fast and understand the whole idea of them. Because I know that in the TOEFL test, we need to be very quick so that we can answer the questions within limited time. (Student49)

In addition to these positive comments, there are also many other negative ones. To begin with, they are of the opinion that as the percentage of readers is very low, they feel unmotivated to read the books. Some others say that almost all the preparatory class students read the books because they had to, not because they wanted to improve their reading skills or vocabulary knowledge. Therefore, for them, there is no relationship between getting higher scores from TOEFL especially from the reading section of it and having reader exams:

For me, readers are just there to help me improve my grades so that I can take the TOEFL test at the end. They are just an opportunity to help students whose grade point average is below 65. I do not think the other students feel motivated to read them. (Student156)

One of the comments made by an upper-intermediate student shows a parallelism with the instructors' comments:

I wonder how many students are really reading the books. Aren't they just listening to their friends who read them? Both the ones who read them and the ones who do not are getting higher grades. Therefore, if you find somebody to tell you the story at the last ten minutes and revise the vocabulary items for just five minutes, I believe you can get more than 70 points. (Student58)

Especially in the first term, they believe that the books chosen for the reader exam are not appropriate for their level. They think that as they are not as challenging as they could be, it makes them feel unmotivated and as a result, they are uneager to read. Therefore, according to one of the B Level student:

Reader exams are quite easy. That's why some of the students get higher grades by just skimming through it or reading the summary. I do not believe there is a relationship between these exams and TOEFL. They might be useful only when very challenging questions are asked or advanced books are chosen to be read. (Student43)

There are two more issues raised related to reader exams. The first one is the length of readers. Even though some of the readers are quite short with at most ten pages, they also have readers which may have forty to fifty pages. Thus, the students find the latter unnecessary and unrelated to their success for the TOEFL test. The second one is about types of questions asked in these exams. They find some of the questions rather detailed even though they think that the books are below their level. Therefore, they believe that they are in a way forced to memorize each and every detail rather than understanding the whole idea. Therefore, they suggest changing both the books and the reader exams in a way that will be more useful for them:

I think the books are not appropriate for our level. They are way too easy. (Student314)

79

Sometimes they are asking very detailed questions which only a limited number of students can answer. I cannot memorize every detail about the book. (Student200)

I do not believe they are effective. First of all, we should not be required to read the long books. Then, books with interesting topics should be chosen and of course they should be appropriate to our level. (Student253)

4.3.3 Comparison of the Results of Instructor and Student Questionnaire

To begin with, when instructors' responses given to three open-ended questions about the effectiveness of three different assessment methods, (midterms, quizzes and readers) are compared to those of the students, it is evident that there are many similarities as well as differences in general. To begin with, there are some issues raised in the answers of both parties. However, the number of instructors who contributed with some suggestions was more than the students. In addition, they make more harsh comments such as "no need to have any speaking quiz" or "any question types rather than multiple choice and error recognition should be omitted from the assessment practices".

For the midterms and quizzes, there were similar comments in both student and instructor questionnaire. Both parties agree on the fact that midterms are harder, covering more subjects, which makes them to be a better practice for the upcoming TOEFL test. The other similarity was related to the last three midterms, which are exact copy of TOEFL ITP, both in terms of the number and the type of questions. Both parties agree on the fact that they are the most effective and useful one; therefore, their predictive power value will be the highest among all the other assessment practices.

For reader exams, there are both instructors and students who agreed on the fact that they can be beneficial for vocabulary development. However, when it comes to the problems related to these assessment practices, their responses were different. Instructors are into suggesting some other ways of assessing what the students have

understood from a piece of reading after talking about the problems such as students' lack of motivation or their reading just the summary of the books. Unlike students, there was not a single instructor to share his/her opinion on the types of questions asked in the exam or the length of the books being chosen. On the other hand, students, instead of finding different ways to improve these exams, commented on the problem. They put the blame on the types of questions, the length and level of the books. They, most of the time, expressed that what makes these reader exams "unsuccessful" was the type of questions asked and the type of book chosen not the fact that they were not reading the books but just making a friend summarize it for them or reading the short summary.

CHAPTER V

DISCUSSION of the RESULTS and CONCLUSION

In this chapter, discussions of the findings, implications for teaching/testing at TOBB ETU as well as suggestions for further research are presented.

5.1 Discussion of the Results

In this study, the relationship between various assessment practices applied in the Department of Foreign Languages at TOBB ETU was examined and the predictive power of these tests was analyzed. Moreover, both the instructors' and students' opinion on the subject was taken. This step helped the researcher to conduct an in-depth analysis on the matter that is being investigated. This part is presented in relation to the research questions of the study.

5.1.1 Research Question 1

Is there a significantly positive relationship among the different types of achievement tests including midterms, quizzes, readers and placement exam grades and TOEFL ITP scores of students studying at the preparatory program at TOBB University?

When the result section of the study is investigated, it can easily be seen that there is a strong positive relationship between the students' exam scores and TOEFL ITP regardless of their level. That is, the students having higher marks from these assessment practices will most probably get higher marks from the TOEFL ITP or vice versa. Moreover, the assessment practices conducted and used by the Testing and Evaluation Office were prepared to check whether the students are ready for the

upcoming test or not. Therefore, they are expected to be a reliable and best indicator of how these students perform in original TOEFL ITP. As it is clear from the analysis of the data, almost all of these assessment practices (midterms, quizzes, readers, placement exam) have a consistently positive relationship with TOEFL ITP across all levels. The only difference was for the placement test: While no significant correlation was found for beginner and elementary levels, a low but significant correlation exist in intermediate and upper-intermediate levels. However, for the other assessment practices especially midterms and quizzes, the results were different. To make it more specific, in Levels A, AF and B, there are a total of ten midterms and twelve quizzes. For the C Level, there are three midterms and three quizzes. That means the relationship between thirty three midterms, thirty nine quizzes and TOEFL ITP is checked. Out of seventy two assessment practices in total, a high correlation is found between all of them except Quiz-1 for AF Level. In other words, 98.62% of the assessment practices prepared at the Testing and Evaluation Unit of the Department of Foreign Languages, TOBB ETU have power in predicting the TOEFL scores of students. This shows that not only each test member responsible for preparing these assessment practices for each level but also the Testing and Evaluation Unit as a whole are successful in designing tests with high predictive power.

Contrary to most students' belief, even between the quizzes which assess students' speaking skills and the official TOEFL ITP test, a strong positive correlation is found. In addition, the types of questions asked in different achievement test conducted at TOBB ETU are very similar to the ones asked in TOEFL ITP. However, there are also some other parts of the assessment practices conducted in the Department of Foreign Languages at TOBB ETU such as the section of vocabulary –a section not included in the official TOEFL test. However, as language is regarded as a whole, with all skills being hand in hand, it is not surprising to see positive relationship between the tests assessing different skills and TOEFL ITP.

83

5.1.2 Research Question 2

How well do the different types of achievement tests including midterms, quizzes, readers and placement exam grades predict the TOEFL ITP scores of students studying at the preparatory program at TOBB University?

The findings of this study suggest that the predictive validity of the assessment practices used in the Department of Foreign Languages at TOBB ETU is high. In other words, the students who have been successful in these tests throughout the term will also get higher scores from the TOEFL at the end. Therefore, these test scores of students can be used to make inferences about students' success for TOEFL ITP. Moreover, they can be useful to diagnose the students with weaknesses. As a follow-up, necessary steps can be taken so that the success of these students might increase. These steps might be opening courses at the weekends or after sessions, or giving private courses to those students. In a similar study, Küçük (2007), it was found out that the predictive validity of the assessment practices prepared and conducted in the Preparatory School of Zonguldak Karaelmas University is high. She suggested encouraging students to take part in the summer courses.

Moreover, the present study yielded the same results from the literature. In this study, the predictive value for different assessment practices range from $r = 26$ to $r = 77$, $p = <.05$, meaning their being moderate to high. Likewise, Ösken (1999) examined the predictive validity of six midterms and a final assessment test conducted in the Department of Basic English at Hacettepe University. Similar results were found as the predictive value of the assessment methods investigated in her study was very high differing from $r = .90$ to $r = .72$.

In addition, last three midterms in all four levels have the highest positive predictive value. That these three midterms are exactly alike TOEFL test can be the explanation of high relationship between them and TOEFL scores. Moreover, this also shows the

importance of test format in predictive validity. In addition, as they are applied at the end of the term, when the whole process of language learning is about to be completed, the students are provided with necessary skills and abilities. In other words, as time passes, they become individuals ready to meet the objectives of the English Preparatory Program. That is why these midterms have the highest predictive value. Likewise, even though there was a positive relationship between both the first quizzes, first midterms and TOEFL ITP, the predictive value was smaller compared to the assessment practices carried out at a later time. This might again be an indication of the effect of time.

5.1.3 Research Question 3

What are the students and instructors perceptions towards the relationship between all these assessment practices and TOEFL ITP?

Among the variables to be examined in this study, quizzes are the first one. In addition to find out the size and direction of the relationship by using statistical procedures, an open-ended question was added to the questionnaire given both to the students and instructors along with a rating scale. By the help of this rating scale, both parties were asked to give a point from one to six to show the power of this relationship between these two assessment practices. The average point that instructors gave is slightly higher than the point given by the students, which means instructors believe that quizzes are more efficient for students' success in TOEFL test. In addition, the answers to the open-ended questions were examined. Here, it is important to clarify that there is a difference between the comments of instructors and students. The instructors think that speaking quizzes are necessary even though it is not directly assessed in the TOEFL ITP. Students, on the other hand, find them completely unnecessary. However, as it is mentioned above, regardless of the levels, a positive relationship is found between these speaking quizzes and students' success of TOEFL ITP even though the magnitude of these relationship may not be as high

as the other quizzes. This shows that all of the components of a language have utmost importance. It is not possible to claim if a specific skill or area is not covered in a language test; it can be avoided as some students suggested. As instructors are more aware of the importance of the fact that language should be considered as a whole rather than its parts separately, many put forward the idea that assessing students for their speaking or writing ability is absolutely necessary as suggested in Ösken (1999) and Özkanal's (1998) studies. However, as in the TOEFL test, these language skills are not evaluated; most students regard them as unnecessary and as a result, they do not give as importance as they give to the other skills. Unfortunately, because of the negative back wash effect of the testing on teaching and learning in the institution where this study was carried out, both the students and instructors may get confused and disregard the importance of writing and speaking skills. Even though it is important for any well-designed test to bring about a desirable wash back, it might not be the case (Hamp- Lyons, 1997). To avoid this, a need for a change is necessary. A test -either prepared by the university itself or the one internationally accepted- which assesses all the skills including writing and speaking might be an option to consider. Ösken (1999) also suggested in her thesis that the assessment test, conducted at the end of the term in the Department of Basic English at Hacettepe University, should be revised. Like the test used in the Department of Foreign Languages at TOBB ETU, there are some skills which are not included such as writing, listening and speaking in theirs, too. Therefore, a revision of the present test or adaptation of a new test might be taken into account by the administration. Likewise, Özkanal (1998) suggested in his study that to increase the validity of the placement test administered in the Foreign Languages Department of Osmangazi University, writing section of it should be changed by either adding paragraph writing or a TOEFL like writing section.

The second variable to be examined in this thesis was midterms. By the help of the rating scale which required both parties to give a point from one to six to show the

power of the relationship between midterms and TOEFL test, the average point was calculated. The point given both by the students and instructors was the highest compared to readers and quizzes. That shows participants find midterms as the most beneficial and effective assessment practice for getting higher scores from TOEFL ITP. More specifically, both the instructors and students compared midterm and quizzes and thought that the former have higher positive relationship in comparison to the latter. For the reason behind it, they mention higher number of questions, similarity between the difficulty level of midterms and TOEFL test, and lastly a wide range of content covered in midterms. With the help of data analysis, similar results were obtained. In addition, in the literature review, similar results were also obtained. In a study carried out by Serpil (2000), instructors' opinion about different assessment methods namely midterms were taken. They put forward that the representation of the course content on the test range from moderate to high. Here, in our study, both the students and the instructors also thought that a wide range of content is covered in the midterms when compared to the quizzes. That is why a higher predictive value between the former and TOEFL ITP was expected by both the students and instructors and the data analysis showed had parallel results.

However, some also mention the fact that the midterms carried out in the first term were not as beneficial as the last ones since they were less difficult and had less similar questions to TOEFL but rather had some productive types such as rewrite or fill-in-blanks. When the relationship between all the midterms and TOEFL ITP is examined, a higher correlation is found between the midterms applied in the last term compared to the ones in the first term. Another similarity is found between the instructors and students comments on the relationship between last three midterms and TOEFL ITP. These last three midterms are exact copies of TOEFL test in terms of number and types of questions, the content, duration and sections. Therefore, both parties thought that the size of relationship will be greater. Data analysis of the students' scores also provided the same results.

The third and last variable to be examined in this thesis was reader exams. After examining the size and direction of the relationship as an answer to first and second research question, the comments of students and instructors to the issue were analyzed for the third research question. Therefore, the average point, which was given from one to six to show the power of the relationship between reader exams and TOEFL test, was found. It was the lowest compared to other two variables, which shows that participants find reader exams as the least beneficial and effective assessment practice for getting higher scores from TOEFL ITP. In addition, various controversial comments are made to this assessment practice. Even though there were a few participants who claim that they might be useful for vocabulary development and reading skills, most claim that reader exam are not as useful as they might be. In this present study, the relationship between these reader exams and the reading part of TOEFL ITP is not checked. However, when the average of students' reader grades and TOEFL ITP scores is examined, a positive relationship is found as mentioned above. On the other hand, most of the participants still think that even though the idea of having reader exams which aim to evaluate students' understanding of a pre-determined book seems to be a useful one, most students just ask a friend to summarize it for them, which makes these exams less efficient. Therefore, instructors demanded that the suggestions proposed by them be considered. Some offered to have different type of an assessment practices rather than a multiple choice written exam such as presentations, role plays or portfolio tasks.

5.2 Implications for Teaching/Testing at TOBB ETU

In order to have accurate results, it is very significant to evaluate assessment practices from various perspectives and different sources. Therefore, even though the predictive value for each test is high, the responses given by students and instructors to open-ended questions provided a basis for suggestions which might be useful for

the Department of Foreign Languages at TOBB ETU. To start with, in TOEFL ITP, there are only three sections which assess students' listening, reading skills along with grammar knowledge. The lack of speaking and writing may put the instructors in dilemma while deciding on which language points to emphasize. If they mistakenly ignore the importance of speaking and writing because of the upcoming test, the failure will be indispensible. Therefore, in order to avoid negative wash back effect, it might be useful to clarify the objectives of the English Preparatory Program of TOBB ETU. In addition, a way to assess all the skills or a test that covers all the areas of a language might be an option to consider.

The most problematic assessment practice seems to be reader exams in the eyes of both the students and instructors. Therefore, Testing and Evaluation Unit might come up with new ways of assessing students' understanding of a reading passage. Some of the suggestions put forward by the instructors might be taken into consideration. Lastly, the results of this study can be used to identify students who are unlikely to get adequate points from the TOEFL test to continue their education in their departments. If the necessary steps are taken to prepare those students better, their changes of being successful can be increased.

5.3 Implications for Further Research
The results of the present study may suggest the followings for future research

1. The study is carried out in the semester of 2012-2013; therefore, it can be replicated in the following years to validate the findings. Furthermore, the results of the TOEFL exam can also be compared with students' scores in the first year departmental English courses. For the success of the preparatory school in the long term, high correlation between the assessment methods and TOEFL as well as students' GPA for their English courses in their departments and TOEFL is expected.

2. In this study, only three different variables are considered and examined. However, there might be many different ones to effect students' scores for TOEFL test. Therefore, future studies may include different variables and examine the relationship between these variables and TOEFL ITP success such as absence or homework. It might be useful to examine the effect of other variables which may also be influential for the success of our students like the one. Therefore, studies carried out by İleri (2013) examining the relationship between students' homework grades and TOEFL ITP or by Roby (2003) examining the relationship between attendance and students' achievement might be useful.

3. This study does not investigate the relationship between the sections of each assessment tests (midterms and quizzes) separately and TOEFL ITP scores. Therefore, a future study might investigate the relationship between the sections of each assessment practices and TOEFL ITP. In this way a more detailed comparison and analysis can be done.

REFERENCES

Al-Ansari S., and Al-Musawi M. (1999). Test of English as a Foreign Language and First Certificate of English tests as predictors of academic success for undergraduate students at the University of Bahrain. *System*, 27(3), 389-99.

Aiken, L. R. (1997). *Psychological testing and assessment.* Boston: Allyn and Bacon.

Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation.* Cambridge: Cambridge University Press.

Bachman, L.F. (1990). *Fundamental considerations in language testing.* New York: Oxford University Press.

Bronner, S., Kennet-Cohen, T. and Oren, C. (1999). *The predictive validity of the components of the process of selection of candidates for higher education in Israel.* National Institute of Testing and Evaluation of Israel.

Brown, J.D. (1996). *Testing in language programs.* New Jersey: Prentice Hall Regents.

Brown, H.D. (2004). *Language assessment: Principles and classroom practices.* New York: Longman.

Brown J. D., Davis J., and Takahashi C. (2012). *Linking, validating and predicting TOEFL IBT scores at advanced proficiency EIKEN levels,* University of Hawai at Manoa Keita Nakamura, Society for Testing English Proficiency.

Carty, H., and Armstrong W. (2008). *The predictive validity of the new SAT achievement test and high school grades on undergraduate success.* University of California, San Diago.

Davies, A. (1990). *Principles of language testing.* Oxford: Basil Blackwell Ltd.

Deno, S. L. (1985). Curriculum-based Measurement: The Emerging Alternative. *Exceptional Children*, *52*(3), 219-232.

Deno, S. L. (2003). Developments in Curriculum-based Measurement. *Journal of Special Education*, *37*(3), 184-192.

Donald G. & Christina P. (2005). *A predictive validity study of the Pharmacy College Admission Test.*

Dooey, P. (1999). *An investigation of the predictive validity of the IELTS test as an indicator of future academic success.* Retrieved from http://cea.curtin.edu.au/tlf/tlf1999/dooey.html.

Enginarlar, H. (2002). *Principles of language testing.* Ankara: METU.

Frankel, J. R., & Wallen, N. E. (2006). *How to design and evaluate research in education.* New York: McGraw Hill.

Güzeller, C. (2005). *The concurrent validity between primary school academic success grades and OKÖSYS subtest grades*. Master Thesis. Akdeniz University.

Hamp-Lyons, L. (1997). Washback, impact, and validity: Ethical concerns. *Language Testing*, *14*(3), 295–303.

Hughes, A. (2003). *Testing for language instructors* (2nd Ed.). Cambridge: Cambridge University Press.

İleri, S. A. (2013). *TOBB ETU preparatory school students' attitudes to homework and the effect of the homework on students' success.* Master Thesis. Ankara University.

Kan, A. (2005). ÖSS'ye kaynaklık eden alan derslerindeki başari ile ÖSS'den elde edilen puanlar arasindaki ilişkinin incelenmesi. *Eğitim ve Bilim. 30*(137), 38-44.

Karakaya I. (2007). Öğretmenlik programlarındaki öğrencilerin ÖSS puanları ile akademik başarıları arasındaki ilişkinin incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 2*(1), 155-163.

Kobrin, J. Patterson, B. Shaw, E. Mattern, K. and Barbuti, S. (2008). *Validity of the SAT for predicting first-year college grade point average,* College Board Research Report, New York: The College Board.

Küçük, F. (2007). *The relationship among face validity, reliability and predictive validity of university EFL Preparatory School achievement tests.* Bilkent University, Master Thesis.

Lado, R. (1961). *Language testing: The construction and use of foreign language tests. A teacher's book.* McGraw-Hill Book Company, New York.

Mattax, D. (1995). *A predictive validity study of the kindergarten diagnostic instrument.* Dissertation. The Ohio State University.

McMullen, S. (2007). *The impact of homework time on academic achievement,* The University of North Carolina at Chapel Hill.

McNamara, T. (2000). *Language testing.* New York: Oxford University Press.

Morgan, R. (1989). An analysis of the predictive validity of the SAT and high school grades from 1976 to 1985. *College Board Research Report, 89* (7) New York: The College Board.

Noble, J. and Sawyer R. (2004). Is high school GPA better than admission test scores for predicting academic success in college? *College and University. 79*(4): 17-22.

Obioma, G., and Salau, M. (2007). The predictive validity of public examinations: A case study of Nigeria. Nigerian Educational Research & Development Council Research Report (NERDC).

Olani, A. (2008). Predicting first year university students' academic success, *Journal of Research in Educational Psychology*, *7*(3), 1053-1072.

Ösken, H. (1999). *An assessment of the validity of the midterm and the end of course assessment tests administered at Hacettepe University, Department of Basic English.* Unpublished Master's Thesis, Bilkent University, Ankara.

Özkanal, Ü. (1998). *An investigation of validity of the placement test administered by Osmangazi University Foreign Languages Department.* Master's Thesis, Bilkent University, Ankara.

Parry, J. M, Mathers, J.M. Stevens, A.J., Parsons, A, Lilford, R, Spurgeon, P, and Thomas, H. (2006). Admissions processes for five year medical courses at English schools: Review. BMJ *332*, 1005-1009.

Roby, E. D. (2004). Research on school attendance and student achievement: A study of Ohio Schools. *Education Research Quarterly*, *28*: 3-4.

Serpil, H. (2000). *An assessment of the content validity of the midterm achievement tests administered at Anadolu University, Foreign Languages Department.* Bilkent University, Master Thesis, Ankara.

Seungsoo Y. (2010). Predicting performance on state achievement tests using Curriculum-Based Measurement in reading: A multilevel Meta-Analysis. *Remedial and Special Education, 31*:412.

Sevindik, H. (2009). *Akademik başarı puanlarının Seviye Belirleme Sınavı (SBS) 2008 puanları ile ilişkisi,* Yüksek Lisans Tezi, Hacettepe Üniversitesi, Ankara.

Shulruf, B, Poole, P, Rudland, J, and Wilkinson, T. (2011). How well do selection tools predict performance later in a medical programme? *Advances in Health, Sciences, Education 17*:615-626.

Shuttleworth, M. (2009). *Predictive Validity.* Experiment Resources, Retrieved Feb 23, 2013 from: http://explorable.com/predictive-validity.

Slavin, R. E. (2007). *Education research in an age of accountability*. Boston, MA: Pearson Education, Inc.

Tokat, E. (2004). *A research relating to the predictive validity of postgraduate education entrance examination and other assent measures.* Master Thesis, Ankara University.

Tokat, E. and Demirtaşlı N. (2004). Les ve diğer kabul ölçülerinin yordama geçerliliğine ilişkin bir çalışma, *Eğitim Bilimleri ve Uygulama*, *3*(5), 35-55.

Yıldırım, A. and Şimşek, H. (1999). *Sosyal bilimlerde nitel araştırma yöntemleri,* (Qualitative research methods in social sciences). Ankara: Seçkin Yayınevi.

Zırhlıoğlu G, and Atlı M. (2011). The predictive validity of the private skill exam scores on academic success in a Physical Education Department, *Education and Science*, *36*: 161-176.

APPENDIX A

TOBB ETÜ'de UYGULANAN SINAVLARIN TOEFL ITP BAŞARISIYLA OLAN İLİŞKİSİ

Sevgili öğrenciler,

TOBB ETÜ'de uygulanan sınavların TOEFL ITP başarısıyla olan ilişkisi hakkında bir yüksek lisans tez çalışması yürütmekteyim. Ekteki anket bu nedenle geliştirilmiştir ve iki bölümden oluşmaktadır. İlk bölüm demografik bilgileri toplamak amacıyla hazırlanmıştır. İkinci bölüm ise konu hakkındaki soruları içermektedir. Bu bölüm için bu soruları yanıtlamanız ve fikirlerinizi paylaşmanız beklenmektedir. Soruların doğru ya da yanlış cevapları yoktur. Cevaplarınızın notlarınız üzerinde herhangi bir etkisi yoktur. Anketten sağlıklı sonuçlar elde edilmesi vereceğiniz samimi cevaplara bağlıdır.

İsim ve soy isminizi yazmanıza gerek yoktur. Vereceğiniz cevaplar sadece araştırma için kullanılacak ve izniniz olmadan başka bir amaç için kullanılmayacaktır.

Katkınız için teşekkür ederiz.   ŞEYMA DOĞRU

e-mail: sdogru@etu.edu.tr

BÖLÜM 1

KİŞİSEL BİLGİLER

1. Bölümünüz _____   **2. Yaşınız** _____

3. Cinsiyetiniz   Bayan ☐   Bay ☐

4.Ne kadar süredir İngilizce öğreniyorsunuz?

1 yıldan daha az ☐   1-3 yıl arası ☐   3-6 yıl arası ☐   6 yıldan fazla ☐

5. Ne tür bir liseden mezun oldunuz?

Kolej   ☐   Düz Lise   ☐   Meslek Lisesi   ☐

Anadolu Lisesi ☐   Fen lisesi   ☐   Anadolu Öğretmen Lisesi ☐

Diğer _____

6.Daha önce hiç İngilizce konuşulan bir ülkede bulundunuz mu?

Evet (7.sorudan devam ediniz) ☐   Hayır (8. Soruya geçiniz) ☐

7. Orada ne kadar süre kaldınız?

1 haftadan daha az ☐   1 hafta-1 ay arası ☐

1 yıla kadar   ☐   1 yıldan daha fazla ☐

8. Ana dili İngilizce olan akraba ya da arkadaşlarınız var mı?

Evet (9.sorudan devam ediniz) ☐   Hayır (11. Soruya geçiniz) ☐

9. Onlarla ne sıklıkla görüşüyorsunuz?

_____

10. Onlarla ne tür bir iletişim aracıyla görüşüyorsunuz?

_____

11. Ana dili İngilizce olan öğretmeniniz oldu mu?

Evet (12.sorudan devam ediniz) ☐   Hayır (13. Soruya geçiniz)☐

12. Ondan haftada kaç saat ne kadar süreyle ders aldınız?

_____

_____

13. Günde ortalama kaç saat İngilizce çalışıyorsunuz?

_____

_____

14. İngilizcenizi geliştirmek adına ne tür aktivitelerle uğraşıyorsunuz?

_____

_____

_____

BÖLÜM 2

15. Sizce hazırlık sınıfında uygulanan kısa sınavların (quizzes) TOEFL ITP başarınıza olan etkisi nedir?

_____

_____

_____

16. 1'den 6'ya kadar bir ölçekte hazırlık sınıfında yapılan bu kısa sınavların (quizzes) TOEFL ITP'deki başarınıza etkisine kaç puan verirsiniz? (1 hiç bir etkisinin olmaması, 6 ise etkisinin en fazla düzeyde olması anlamına gelmektedir, seçmek istediğiniz puanı lütfen yuvarlak içine alınız.)

1	2	3	4	5	6

17. Sizce hazırlık sınıfında uygulanan ara sınavların (midterms) TOEFL ITP başarınıza olan etkisi nedir?

_____
_____
_____

18. 1'den 6'ya kadar bir ölçekte hazırlık sınıfında yapılan bu ara sınavların (midterms) TOEFL ITP'deki başarınıza etkisine kaç puan verirsiniz? (1 hiç bir etkisinin olmaması, 6 ise etkisinin en fazla düzeyde olması anlamına gelmektedir, seçmek istediğiniz puanı lütfen yuvarlak içine alınız.)

1	2	3	4	5	6

19. Sizce hazırlık sınıfında uygulanan okuma sınavlarının (readers) TOEFL ITP başarınıza olan etkisi nedir?

_____
_____
_____

20. 1'den 6'ya kadar bir ölçekte hazırlık sınıfında yapılan bu okuma sınavlarının (readers) TOEFL ITP'deki başarınıza etkisine kaç puan verirsiniz? (1 hiç bir etkisinin olmaması, 6 ise etkisinin en fazla düzeyde olması anlamına gelmektedir, seçmek istediğiniz puanı lütfen yuvarlak içine alınız.)

1	2	3	4	5	6

TOBB ETÜ'de UYGULANAN SINAVLARIN TOEFL ITP BAŞARISIYLA OLAN İLİŞKİSİ

Sevgili okutmanlar,

TOBB ETÜ'de uygulanan sınavların TOEFL ITP başarısıyla olan ilişkisi hakkında bir yüksek lisans tez çalışması yürütmekteyim. Ekteki anket bu nedenle geliştirilmiştir ve iki bölümden oluşmaktadır. İlk bölüm demografik bilgileri toplamak amacıyla hazırlanmıştır. İkinci bölüm ise konu hakkındaki soruları içermektedir. Bu bölüm için bu soruları yanıtlamanız ve fikirlerinizi paylaşmanız beklenmektedir. Soruların doğru ya da yanlış cevapları yoktur. Anketten sağlıklı sonuçlar elde edilmesi vereceğiniz samimi cevaplara bağlıdır.

İsim ve soy isminizi yazmanıza gerek yoktur. Yanıtlarınız izniniz olmadan başka bir amaç için kullanılmayacaktır.

Katkınız için teşekkür ederiz.   ŞEYMA DOĞRU
                                                   e-mail: sdogru@etu.edu.tr

BÖLÜM 1

KİŞİSEL BİLGİLER

1.Ne kadar süredir İngilizce öğretiyorsunuz?

1 yıldan daha az ☐   1-3 yıl arası ☐   3-6 yıl arası ☐   6 yıldan fazla ☐

2.Ne kadar süredir TOBB ETÜ'de çalışıyorsunuz?

1 yıldan daha az ☐   1-3 yıl arası ☐   3-6 yıl arası ☐   6 yıldan fazla ☐

3.Yaşınız _____   **4. Cinsiyetiniz**   Bayan ☐   Bay ☐

5. Daha önce hiç İngilizce konuşulan bir ülkede bulundunuz mu?

Evet (6.sorudan devam ediniz) ☐   Hayır (İkinci bölüme geçiniz) ☐

6. Orada ne kadar süre kaldınız?

1 haftadan daha az ☐   1 hafta-1 ay arası ☐

1 yıla kadar ☐   1 yıldan daha fazla ☐

BÖLÜM 2

7. Sizce hazırlık sınıfında uygulanan kısa sınavların (quizzes) öğrencilerimizin TOEFL ITP başarısına olan etkisi nedir?

_____
_____
_____

8. 1'den 6'ya kadar bir ölçekte hazırlık sınıfında yapılan bu kısa sınavların (quizzes) öğrencilerimizin TOEFL ITP başarısına olan etkisine kaç puan verirsiniz? (1 hiç bir etkisinin olmaması, 6 ise etkisinin en fazla düzeyde olması anlamına gelmektedir, seçmek istediğiniz puanı lütfen yuvarlak içine alınız.)

1	2	3	4	5	6

9. Sizce hazırlık sınıfında uygulanan ara sınavların (midterms) öğrencilerimizin TOEFL ITP başarısına olan etkisi nedir?

_____
_____
_____

10. 1'den 6'ya kadar bir ölçekte hazırlık sınıfında yapılan bu ara sınavların (midterms) öğrencilerimizin TOEFL ITP başarısına olan etkisine kaç puan verirsiniz? (1 hiç bir etkisinin olmaması, 6 ise etkisinin en fazla düzeyde olması anlamına gelmektedir, seçmek istediğiniz puanı lütfen yuvarlak içine alınız.)

1	2	3	4	5	6

11. Sizce hazırlık sınıfında uygulanan okuma sınavlarının (readers) öğrencilerimizin TOEFL ITP başarısına olan etkisi nedir?

_____
_____
_____

12. 1'den 6'ya kadar bir ölçekte hazırlık sınıfında yapılan bu okuma sınavlarının (readers) öğrencilerimizin TOEFL ITP başarısına olan etkisine kaç puan verirsiniz? (1 hiç bir etkisinin olmaması, 6 ise etkisinin en fazla düzeyde olması anlamına gelmektedir, seçmek istediğiniz puanı lütfen yuvarlak içine alınız.)

1	2	3	4	5	6

**THE QUESTIONNAIRE ABOUT THE RELATIONSHIP BETWEEN THE
ASSESSMENT PRACTICES CARRIED OUT AT TOBB ETÜ AND TOEFL
ITP**

Dear Students,
I am carrying out a study to investigate the effectiveness of reading room and self-
access center. The attached questionnaire has been developed for this purpose and it
consists of two sections. The first section asks for some demographic information. In
the second section, there are some questions related to your experience or the time
that you spend in either of these places. For this second section, you are expected to
go over the questions and write the related information and your opinions. There is
no right or wrong answers to the questions. You <u>will not be graded</u> based on your
answers. Answering all the questions is important for the study to be carried out
completely. Your sincere responses to the questions are highly important to get
accurate results.

You do not have to write your name and surname. You responses to the questions
will not be used for any other purpose without your permission.

Thank you for your cooperation.
Regards,   ŞEYMA DOĞRU
                                                       e-mail: sdogru@etu.edu.tr

SECTION I

PERSONAL INFORMATION

1. Your Major _____   **2. Your age** _____

3. Your gender   Female ☐   Male ☐

4. How long have you been learning English?

Less than 1 year ☐   1-3 years ☐   3-6 years ☐   more than 6 years ☐

5. What kind of high school did you graduate from?

College ☐   Regular High School ☐   Vocational High School ☐

Anatolian High School ☐ Science High School   ☐

Anatolian Teacher Training High S. ☐   Other _____

6. Have ever been in a country where English is the official language?

Yes (Continue with the question 7) ☐   No (Pass to the question 8) ☐

7. How long did you stay there?

Less than 1 week☐   1week-1month☐ 1 year ☐   more than 1 year ☐

8. Do you have any relatives or friends whose native language is English?

Yes (Continue with the question 9) ☐   No (Pass to the question 11) ☐

9. How often do you get in touch with them?

_____

10. How do you get in touch with them?

_____

11. Did you have any teachers whose native language was English?

Yes (Continue with the question 12) ☐   No (Pass to the question 13)   ☐

12. How many hours in a week were you lectured by him or her and for how long?

_____

_____

13. On average how much time do you spend in a day to study English?

_____

_____

14. What kinds of activities do you engage with?

_____

_____

SECTION 2

15. In your opinion, what is the effect of quizzes on your success of TOEFL ITP?

_____
_____
_____

16. On a scale of 1 (is not effective at all) to 6 (is highly effective) what point would you give to the effect of quizzes on your success of TOEFL ITP?

1	2	3	4	5	6

17. In your opinion, what is the effect of midterms on your success of TOEFL ITP?

_____
_____
_____

18. On a scale of 1 (is not effective at all) to 6 (is highly effective) what point would you give to the effect of midterms on your success of TOEFL ITP?

1	2	3	4	5	6

19. In your opinion, what is the effect of reader exams on your success of TOEFL ITP?

_____
_____
_____

20. On a scale of 1 (is not effective at all) to 6 (is highly effective) what point would you give to the effect of reader exams on your success of TOEFL ITP?

1	2	3	4	5	6

THE QUESTIONNAIRE ABOUT THE RELATIONSHIP BETWEEN THE ASSESSMENT PRACTICES CARRIED OUT AT TOBB ETÜ AND TOEFL ITP

Dear Instructors,

I am carrying out a study to investigate the effectiveness of reading room and self-access center. The attached questionnaire has been developed for this purpose and it consists of two sections. The first section asks for some demographic information. In the second section, there are some questions related to your experience or the time that you spend in either of these places. For this second section, you are expected to go over the questions and write the related information and your opinions. There is no right or wrong answers to the questions. Answering all the questions is important for the study to be carried out completely. Your sincere responses to the questions are highly important to get accurate results.

You do not have to write your name and surname. You responses to the questions will not be used for any other purpose without your permission.

Thank you for your cooperation.

Regards,   ŞEYMA DOĞRU

                                                                e-mail: sdogru@etu.edu.tr

SECTION I

PERSONAL INFORMATION

1. How long have you been learning English?

Less than 1 year ☐   1-3 years ☐   3-6 years ☐   more than 6 years ☐

2. How long have you been working at TOBB ETÜ?

Less than 1 year ☐   1-3 years ☐   3-6 years ☐   more than 6 years ☐

3.Age _____   **4. Gender**   Female ☐   Male ☐

5. Have you ever been in a country where English is the official language?

Yes (Continue with question six) ☐   No (Pass section two) ☐

6. How long did you stay there?

Less than 1 week☐   1week-1month☐   1 year ☐   more than 1 year ☐

SECTION 2

13. In your opinion, what is the effect of quizzes on your success of TOEFL ITP?

_____

_____

_____

14. On a scale of 1 (is not effective at all) to 6 (is highly effective) what point would you give to the effect of quizzes on your success of TOEFL ITP?

1	2	3	4	5	6

15. In your opinion, what is the effect of midterms on your success of TOEFL ITP?

_____

_____

_____

16. On a scale of 1 (is not effective at all) to 6 (is highly effective) what point would you give to the effect of midterms on your success of TOEFL ITP?

1	2	3	4	5	6

17. In your opinion, what is the effect of reader exams on your success of TOEFL ITP?

_____

_____

_____

18. On a scale of 1 (is not effective at all) to 6 (is highly effective) what point would you give to the effect of reader exams on your success of TOEFL ITP?

1	2	3	4	5	6

TEZ FOTOKOPİSİ İZİN FORMU

ENSTİTÜ

Fen Bilimleri Enstitüsü ☐

Sosyal Bilimler Enstitüsü ☐

Uygulamalı Matematik Enstitüsü ☐

Enformatik Enstitüsü ☐

Deniz Bilimleri Enstitüsü ☐

YAZARIN

Soyadı : DOĞRU
Adı   : ŞEYMA
Bölümü : İNGİLİZ DİLİ EĞİTİMİ

TEZİN ADI (İngilizce) : HOW WELL DO VARIOUS ASSESSMENT
PRACTICES OF THE PREPARATORY YEAR ENGLISH PROGRAM PREDICT
THE SUCCESS OF STUDENTS FOR TOEFL ITP (INSTITUTIONAL TESTING
PROGRAM)?

A CASE STUDY AT TOBB UNIVERSITY OF ECONOMICS AND
TECHONOLOGY DEPARTMENT OF FOREIGN LANGUAGES

TEZİN TÜRÜ : Yüksek Lisans ☐   Doktora ☐

1. Tezimin tamamından kaynak gösterilmek şartıyla fotokopi alınabilir. ☐

2. Tezimin içindekiler sayfası, özet, indeks sayfalarından ve/veya bir
bölümünden kaynak gösterilmek şartıyla fotokopi alınabilir. ☐

3. Tezimden bir bir (1) yıl süreyle fotokopi alınamaz. ☐

TEZİN KÜTÜPHANEYE TESLİM TARİHİ: