

**An Investigation of the Predictive Validity of ELT  
Improvement Courses' Exams in an ELT  
Undergraduate Program**

**Yaseen Zeedan**

Submitted to the  
Institute of Graduate Studies and Research  
in partial fulfillment of the requirements for the degree of

Master of Arts  
in  
English Language Teaching

Eastern Mediterranean University  
January 2019  
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

---

Assoc. Prof. Dr. Ali Hakan Ulusoy  
Acting Director

I certify that this thesis satisfies the requirements as a thesis for the degree of Master of Arts in English Language Teaching.

---

Assoc. Prof. Dr. Javanshir Shibliyev  
Chair, Department of Foreign Language Education

We 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 in English Language Teaching.

---

Asst. Prof. Dr. İlkey Gilanlıođlu  
Supervisor

Examining Committee

1. Assoc. Prof. Dr. Javanshir Shibliyev

---

2. Assoc. Prof. Dr. Oytun Sözüdođru

---

3. Asst. Prof. Dr. İlkey Gilanlıođlu

---

## ABSTRACT

This study was designed to investigate the predictive validity of the exams of the English language improvement courses in the undergraduate program as well as the relationship between the scores of these exams. A quantitative methodology was used in this study. The study was conducted at the Foreign Language Education Department (FLE) of the Faculty of Education at Eastern Mediterranean University (EMU). The Data was analyzed using quantitative statistics and presented using tables. The mid-term and final exam scores of 1716 students in the academic years 2013-2018 of FLE, and their final grades in the English improvement courses were correlated.

The results of this study revealed that the bulk of the exams in the improvement courses have strong predictive validity. Moreover, the strong correlations between the exam scores revealed that the exams scores could predict the performance of the students in future exams and tests, and that the majority of these scores represent a real performance of the students. The outcomes revealed that there is positive relationship between the scores of mid-term and the final exam scores; it also proved that the relationship between the final exams scores and the total grades in those courses is strong and positive.

**Keywords:** Predictive Validity, English Language Teaching, Undergraduate, Exams, Improvement Courses.

## ÖZ

Bu çalışma, lisans programındaki İngilizce dil öğretimi iyileştirme kursu sınavlarının yordama geçerliliğini ve bu sınavların puanları arasındaki ilişkiyi araştırmak amacıyla hazırlanmıştır. Bu çalışmanın yürütülmesinde nicel yöntem kullanılmıştır. Çalışma Doğu Akdeniz Üniversitesi Eğitim Fakültesi Yabancı Diller Eğitimi Bölümünde yürütülmüştür. Veriler nicel istatistik yöntemiyle analiz edilmiş ve tablo kullanılarak verilmiştir. 2013-2018 yılları arasında Doğu Akdeniz Üniversitesi Yabancı Diller Eğitimi Bölümünde kayıtlı 1716 öğrencinin ara sınav ve final sınav notları ile İngilizce takviye derslerinde elde ettikleri başarı notları karşılaştırılmıştır.

Çalışmada elde edilen bulgular, iyileştirme kurslarındaki sınavların çoğunun güçlü yordama geçerliliğine sahip olduğunu ortaya koymuştur. Ayrıca sınav puanları arasındaki güçlü korelasyonlar, sınav puanlarının öğrencilerin test ve sınavlarda performansını yordayabildiğini ve bu puanların öğrencilerin gerçek performansını büyük ölçüde yansıttığını ortaya koymuştur. Sonuçlar, ara sınav puanları ile final sınavı puanları arasında pozitif bir ilişki olduğunu ve final sınav puanları ile genel başarı puanları arasında pozitif bir ilişki olduğunu ortaya koymuştur.

**Anahtar Kelimeler:** yordama geçerliliği, İngiliz Dili Eğitimi, lisans, sınav, takviye dersler.

**TO**

My dearest and beloved mother and father

My loving brother and sister

My grandmothers and grandfathers

All my family members and relatives

## ACKNOWLEDGMENT

Having thanked my God. It is my exquisite delight to express my sincere gratitude to my supervisor help. Prof. Dr. İlkey Gilanlıođlu for his outstanding encouragement and help, without his help this work could not occur. He reinforced me with his extraordinary thoughts and helped me in preparing this research. His persistence is undisputable from the primary phases of this study until the finale stage. I am appreciative to him, his support made it possible to achieve this goal.

Furthermore, I would love to record my gratitude to my thesis defense jury members for their useful remarks and their valuable directions. I would also love to thank the employees, the workers, and the students as well as all professors of the Department of Foreign Language Education. Especially, people who educated me during my study years: Assoc. Prof. Dr. Javanshir Shibliyev, Assoc. Prof. Dr. Naciye Kunt, Prof. Dr. Ülker Vancı Osam, and Asst. Prof. Dr. Fatoş Erozan for their fruitful guidance and productive education that I applied while I was composing and writing this thesis.

Eventually, it would be my delight to express my gratefulness to my family members and my friends, my dear father and dear mother for their assist and their empowering words every now and then, which made me, study enough to conduct this work successfully and effectively.

# TABLE OF CONTENTS

ABSTRACT .....	iii
ÖZ.....	iv
DEDICATION.....	v
ACKNOWLEDGMENT.....	vi
LIST OF TABLES.....	x
LIST OF FIGURES.....	xii
LIST OF ABBREVIATIONS .....	xiii
1 INTRODUCTION.....	1
1.1 Background of the Study.....	1
1.2 The Statement of the Problem .....	3
1.3 Purpose of the Study.....	4
1.4 Significance of the Study.....	4
1.5 Research Questions.....	5
2 LITERATUREREVIEW.....	6
2.1 The Meaning of Test Validity.....	6
2. 2 Predictive Validity in English Language Tests.....	8
2.2.1 Standardized and International English Exams.....	9
2.2.2 Locally Developed English Exams Conducted to Test the Relationship between Academic Achievement and Language Performance.....	13
3 METHODOLOGY.....	19
3.1 Research Design.....	19
3.2 Research Context.....	20

3.3	Participants.....	21
3.4	Research Questions.....	21
3.5	Data Collection Instruments.....	22
3.6	Data Collection Procedures.....	22
3.7	Data Analysis Instruments .....	22
3.8	Data Analysis Procedures.....	23
4	RESULTS.....	25
4.1	Descriptive Statistical Analysis for the Students' Exams Scores of Each Course for Each Academic Year .....	25
4.2	The Correlations between the Students' Exam Scores for Each Academic Year.....	31
4.3	Descriptive Statistical Analysis for the Students' Exams Scores of Each Course in the Five Academic Years Together .....	35
4.4	The Correlations between the Students' Exam Scores for all the Five Academic Years.....	40
5	DISCUSSION AND CONCLUSION.....	41
5.1	Discussion of Results.....	41
5.1.1	Question 1: Do the Improvement Courses' Exams Have Predictive Validity.....	41
5.1.2	Question 2: Is There a Connection between the Mid-term Exams' Scores and the Final Exams' Scores? .....	45
5.1.3	Question 3: Is there any Relationship between the Final Exams' Scores and the Courses' Total Grades?.....	45
5.2	Conclusion.....	46
5.3	Implications for the Study.....	46



5.4 Limitations .....	47
5.5 Suggestions for Further Research .....	47
REFERENCES.....	48
APPENDICES.....	55
Appendix A: Approval letter from the Ethical Committee (BAYEK).....	56
Appendix B: Permission Letter from the Department of Foreign Language Education.....	57
Appendix C: Data Collection Permission Letters.....	58
Appendix D: Exams' Scores Distributions Histograms.....	59

## LIST OF TABLES

Table 1: Descriptive Statistics for English Improvement Courses' Exams Scores for the Academic Year 2013-14 .....	26
Table 2: Descriptive Statistics for English Improvement Courses' Exams Scores for the Academic Year 2014-15 .....	27
Table 3: Descriptive Statistics for English Improvement Courses' Exams Scores for the Academic Year 2015-16.....	28
Table 4: Descriptive Statistics for English Improvement Courses' Exams Scores for the Academic Year 2016-17 .....	29
Table 5: Descriptive statistics for English Improvement Courses' Exams Scores for the Academic Year 2017-18.....	30
Table 6: Correlation Coefficient (r) between English Improvement Courses' Exams Scores for the Academic Year 2013-14.....	31
Table 7: Correlation Coefficient (r) between English Improvement Courses' Exams Scores for the Academic Year 2014-15.....	32
Table 8: Correlation Coefficient (r) between English Improvement Courses' Exams Scores for the Academic Year 2015-16.....	33
Table 9: Correlation Coefficient (r) between English Improvement Courses' Exams Scores for the Academic Year 2016-17.....	33
Table 10: Correlation Coefficient (r) between English Improvement Courses' Exams Scores for the Academic year 2017-18.....	34
Table 11: Descriptive Statistics for English Improvement Courses' Exams Scores for the Academic Years 2013-2018.....	36

Table 12: Correlation Coefficient (r) in English Improvement Courses' Exams  
Scores for the Academic Year 2013-18.....40

## LIST OF FIGURES

Figure 1: ELTE 107 mid-terms' exam scores distribution.....	38
Figure 2: ELTE 105 mid-terms' exam scores distribution.....	39

## **LIST OF ABBREVIATIONS**

ELTE 101	Contextual Grammar - I
ELTE 102	Contextual Grammars - II
ELTE 103	Advanced Reading and Writing - I
ELTE 104	Advanced Reading and Writing - II
ELTE 105	Listening and Pronunciation - I
ELTE 106	Listening and Pronunciation - II
ELTE 107	Oral Communication Skills - I
ELTE 108	Oral Communication Skills - II
ELTE 112	Vocabulary Course

# Chapter 1

## INTRODUCTION

This chapter presents the background of the study, the statement of the problem, the purpose of the study, the research questions, and the significance of the study.

### **1.1 The Background of the Study**

Testing has the foremost essential role within the language teaching process because it is the tool to measure the outcomes of the teaching process and the student's level. Given the importance of tests, they should be valid, reliable, and authentic to give accurate results that represent the level and the knowledge of the students. Thus, this study investigates the predictive validity of English exams used in improvement courses in the undergraduate program. Validity refers to how well a test measures what it is supposed to measure (Carmines, & Zeller, 1979; Golafshani, 2003). There are many types of validity such as, criterion validity, content validity, construct validity, face validity (Carmines, & Zeller, 1979). This study will focus on predictive validity, which is a type of criterion validity. An exam was believed to be either valid or not as proved via the connections between the test and some other external criterion measures (Goodwin & Leech, 2003).

In spite of the fact that testing is at present utilized for some reasons in the educational frameworks, Goodwin & Leech, (2003) stated that their viability and utility should eventually be made based on the degree at which they advance student learning. The point of testing ought to be "to teach and enhance students

performance, not only to review it" (Wiggins, 1998, p.7). Teachers should get critical and helpful data from each test. In education, valid decision making relies upon access to important, precise, and opportune data. Moreover, the data picked up ought to be put to great use to enhance students learning (Falk, 2000; National Council of Teachers of Mathematics, 1995).

It is established that testing is the most applicable way to evaluate and systematically rank the students, although some have argued whether testing is necessary or not (Graves, 2002). Within the field of English language teaching and education in general, all experts and scholars highlighted on the necessity of exams and testing, thus, tests remain the most popular way to grade learners, the quality of their production would seem vital. Therefore, exams should be accurate and valid so the results of these exams can represent the actual ability and performance of the students.

Validity and reliability are two of the most significant conditions for test efficiency. Scholars argued that we could achieve accurate evaluation for the students if they are both consistent and guaranteed in a test (Charney, 1984; Clark & Watson, 1995).

Predictive validity, which this study focuses on, involves the assessment of some or all aspects of the subject ability of individuals in some context and for some set of purposes (not necessarily common to all parties). Hence, the importance of predictive validity for examinations cannot be neglected. Predictive validity of tests enhances the value of the tests to students and other stakeholders in the education field. It is thus, essential to understand the predictive validity of internal examinations in English improvement courses in the undergraduate program.

Correlation is a statistical technique to measure a potential direct and linear connotation between two variables (Crowder, 2017). It is simple both to compute and to deduce. Correlation is a process of calculating a potential linear and direct connotation between two factors. Correlation is determined by a measurement called the correlation coefficient, which denotes the strength of the accepted linear connotation between the variables. A dimensionless number takes a significance in the range  $-1$  to  $+1$ . A correlation coefficient of zero shows that no linear relationship occurs among two variables, and a correlation coefficient of  $-1$  or  $+1$  designates a perfect linear relationship (Mukaka, 2012).

## **1.2 The Statement of the Problem**

Creating a good and accurate exam has been one of the major topics in English language field. Scholars have attempted to find the best ways to guarantee and measure the quality of an exam from many aspects. The lack of studies that focused on the validity of teacher-made exams in undergraduate programs, as well as the lack of interest in the quality of exams such as mid-terms and finals exams, in the English improvement courses in the undergraduate programs in universities necessitated the call for this study.

There is also, a problem that many exams in undergraduate programs in universities do not reflect the actual performance of the students, due to several factors. Despite the importance of such exams in the evaluation and the expectation of the students' performance in the rest of the undergraduate program and their ability in English language in general, no studies have tried or showed interested in measuring and evaluating midterm and final exams in English improvement courses in the universities. The importance of such exams necessitated us to evaluate and measure



the validity of such exams as these exams reflect the success and the strength of the undergraduate university programs.

### **1.3 The Purpose of the Study**

This study investigates the predictive validity of the English language teaching improvement courses' exams (mid-terms and finals) in the undergraduate program as well as the relationship between the scores of these exams. In addition, the study also measures the correlation between the mid-term, final and total grades' scores to estimate the relationship between these scores and to know how strong this relationship is. Thus, knowing the predictive power of these exams.

### **1.4 Significance of the Study**

This study will be useful in evaluating the exams of the improvement courses in the Department of Foreign Language Education offered in the undergraduate program at the Eastern Mediterranean University. The study will contribute to the improvement of the exams' program also by finding the relationship between the results of the students in the mid-term and final exams and their final grades in the courses. Furthermore, this study will find the relationship and the meaning of this relationship between the scores of these exams, this relationship will show if these exams have predictive power and validity or not. Thus, giving an idea of whether students are on the right track, and if these exams are valid in predicting the future performance of the students or not. Most of the studies that investigated the predictive validity were based on the GPAs as measures of students' performance. Most studies also did not pay attention to the measurement of the predictive validity of the achievement exams in the English language courses offered in the university, which this study tries to do.

## **1.5 The Research Questions**

This study seeks to provide answers for these questions:

1. Do the exams in improvement courses have predictive validity?
2. Is there a connection between the mid-term exams' scores and the final exams' scores in the English improvement courses?
3. Is there any relationship between the final exams' scores and the courses' total grades?

## **Chapter 2**

### **LITERATURE REVIEW**

This chapter first explains the meaning of test's validity in general. Secondly, it explores predictive validity in English language tests in both standardized and International English exams as well as locally developed English exams, which are conducted to test the relationship between academic achievement and language performance. In addition, this study presents a summary of research studies about the correlation between English language tests scores and future performance of the students.

#### **2.1 The Meaning of Test's Validity**

The notion "validity" has been referred to as "the degree to which empirical evidence and theoretical justifications assist the suitability and the adequacy of interpretations based totally on test scores" (Messick, 1989, p. 13). Notwithstanding the fact that this definition is dated (nearly twenty-five years old), there is a common recognition in the literature for this definition. Thus, evaluation is about test scores, and the interpretations of these scores.

An exam is considered valid if it measures what it supposed to measure indicated by Hughes (2003). Scholars have approached the notion of validity from various magnitudes, for each of which demonstrate a distinguish view to the conception of validity. Hughes (2003) classified validity into four central categories, which are Construct validity, content validity, and Criterion-related validity (predictive as well

as concurrent validity), Face validity. Hughes (1989) and Bachman (1990) gave the following four categories: construct validity, content validity (including external and internal validity), criterion-based validity, and face validity.

Construct validity refers to the degree of accuracy of a construct, which an exam is assumed to compute (Brown 1994; Bachman & Palmer 1996) also, “must exhibit that the tests that the scholars are using are evocative to the applicants themselves” (Cohen et al., 2000 p. 110).

Content validity is troubled with the level to which the items of an exam have to do with the real-life condition they are endeavoring to measure (Bachman 1990; Hughes 1989). In the field of content validity, there are external validity and internal validity, which is concerned with the association between dependent and independent variables when experimentations are carried out. External validity arises when outcomes can be connected to the overall populous, while internal validity is linked to the removal of problematic variables in studies.

Criterion-related validity “relays the findings of one specific tool or a measurement to an external criterion” (Cohen et al. 2000 p. 111). It encompasses two major categories, these being concurrent and predictive validity. Concurrent validity is not required to be calculated over a span of time and can be “demonstrated instantaneously with a different instrument”. Regarding predictive validity, if results from two detached but related experimentations or tests produce similar outcomes the original examinations are said to have strong predictive validity (2000 p. 112).

Face validity relates to what degree an exam is measuring what it is supposed to measure. In general, face validity in testing describes the look of the test as opposed to whether the test is valid or not.

Predictive validity according to Alderson, Clapham and Wall (1995), could be observed for the same test-takers by associating a score of an exam with another measure, such as another exam score, that is gathered after the exam has been set. It is vital to search for predictive validity in a proficiency exam since predictive validity exploration are essential in proving whether the focal purpose of the proficiency exam, which is to assess and measure the test-takers' ability in performing effectively and successfully in a forthcoming course, is attained or not. The time element in predictive validity is important to measure the capability of the test in predicting the performance of the students in the future. This capability of predicting the students' performance in future reflects how strong and valid the test is, which gives more validity to the results and the performance of the students.

## **2.2 Predictive Validity in English Language Tests**

Predictive validity concerns with how accurately a test or an exam score predicts the success and the performance in the future (Weir, 2005). For instance, the scores on an English improvement course test, whose purpose is to evaluate students' level in English skills should highly correlate with the academic and scholastic performance of the same students as measured by their grades on successive courses or other test scores in the same course (Davies et al., 1999). A number of predictive validity researches, both of standardized and international tests or locally developed tests examined the relationship between language performance and academic achievement.

### **2.2.1 Standardized and International English Exams**

Tests such as TOEFL and IELTS, which are well-known exams with well-established production procedure, have specially made both validation studies in an endeavor to prove the claims that they give a sufficient measure of candidate's ability to start training and studying in the environment of English language (IELTS, 2012). In addition to their capability to communicate as well as practice the English language at the academia level (TOEFL iBT Test, 2013).

Dooley (1999) conducted a study to find out if IELTS is a precise predictor of achievement and performance for Engineering, Science and Business undergraduates. The study involved 89 first year students. The findings reflected that positive relationship was found between achievement in business course and IELTS reading subcategory scores.

Huong (2001) conducted a research to examine the IELTS scores predictive validity. The association between other academic performance and IELTS scores was investigated. Huong observed a positive and important correlation between second and first semester GPA's and IELTS scores, which was regarded adequate for Alderson, Clapham and Wall (1995). Furthermore, amongst all the IELTS sub-tests (Writing, Reading, Speaking and Listening) the maximum correlation was detected between the first semester GPA and the reading sub-test; Listening and Reading related to Speaking and Writing had greater correlations with the second and first semester GPA's. The discoveries proposed a correspondence with the GPA of second and first semester in relations with both sub-test scores and the total scores of the IELTS exam.

The researches of language proficiency predictive validity as measured by IELTS have created questionable outcomes. In most researches, the standard criterion of choice regarding the measure of predictive validity was the Grade Point Average (GPA); however, some utilize the GPA of first year or the GPA of first semester to minimize the impact of the moderator variable (Cope, 2011). More or less researchers reported noteworthy weak correlation but positive, between GPA and IELTS scores (Feast, 2002), or modest correlation between first semester GPA and overall IELTS scores (Woodrow, 2006; Yen & Kuzma, 2009). Nevertheless, more or less researchers discovered that the positive correlations is not exist, between either GPA and IELTS scores (Cotton & Conrow, 1998), or diminutive proof for the IELTS exam validity as an academic performance predictor. Ingram and Bayliss (2007) and Dooley & Oliver (2002), on the other hand, examined the connection between three dependent elements: student's insights of their language proficiency, language behavior in their academic courses, staff observations and researcher ratings of the students and their IELTS marks. They discovered that the marks of the IELTS test precisely predicted the language behavior of the students in the four skills in the opening six months of university education. The degree of association between the marks of the IELTS exam and the criterion has been investigated by the studies of predictive validity as well. Conrow and Cotton (1998) showed that reading as well as writing scores correlated directly with the ratings of the staff, while Kerstjens and Nery's (2000) research showed that merely the scores of reading revealed a positive substantial correlation with the GPA of the first-semester. Furthermore, Yen and Kuzma (2009) proved the momentous correlations between the writing, listening and reading scores and the GPA of the first-semester amongst an assembly of students in China. Nonetheless, Woodrow (2006) noted positive correlations among the scores

and the GPA of the first-semester except reading. (Woodrow, 2006) stated that the level of proficiency in English language, influenced the success of the students who score 6.5 or below, while with students who score seven and above, English proficiency has no impact on academic performance.

TOEFL likewise has been a topic for predictive validity researches. Bridgeman and Cho (2012) discovered insignificant correlation coefficients of predictive validity between the GPA and TOEFL iBT with average correlations equals  $r=0.16$  for graduates and  $r=0.18$  for undergraduate students in four topics arts and humanities, business, social sciences and science and engineering. Nevertheless, they also revealed that students in the top 25% GPA group were in the top 25% in TOEFL iBT, as the students in the bottom 25% TOEFL group were.

Even though IELTS and TOEFL are recognized tests of English language academic usage, but some universities in English-speaking countries recognize some other English language certificates and exams. A paper by Grote, Oliver, and Vanderford (2012) examined the predictive validity of 25 kinds of tests of proficiency proof recognized by the universities in Australia. They concluded that even though the exams of English language proficiency had small predictive validity generally, standardized exams such as TOEFL and IELTS are the best pointers of academic success in future.

Some investigations on the association between TOEFL exam and further English proficiency exams was measured, in Marvin and Simner (1999) in order to vindicate the usage of the TOEFL exam for decision-making. Moreover, according to Marvin and Simner (1999) a connection between the accomplishment in the first year of



university English disciplines and TOEFL scores can be potential; nevertheless, the connection may not endure further than the first year. Pack (1972 cited in Marvin & Simner, 1999) conducted a research on 402 learners. He discovered that, the scores of TOEFL exam were "significantly connected to the mark acquired in the English course which was taken first, nevertheless, they are not interrelated to marks achieved in successive English courses nor are they to the possibility of the graduation of the examinee" (Hale et al. p. 161). The Pearson correlations between the Weighted Average Marks (WAM) and the IELTS scores were small. For the undergraduates, the one noteworthy connection was between WAM and the reading IELTS sub-test ( $r=0.27$ ), whereas for the postgraduates, there was a fragile but important connection between the total score and WAM and three subtest marks with the exclusion of writing (the correlations spread from  $r=0.16$  for speaking to  $r=0.29$  for reading).

The undefined outcomes might be due to various causes. Firstly, numerous researches are established on insignificant sample sizes such as 113 (Kerstjens & Nery, 2000), 33 (Cotton, F., & Conrow, 1995), 101 (Feast, 2002), 65 (Dooley & Oliver, 2002), 61 (Yen & Kuzma, 2009). Secondly, the standard in each research is diverse from the other researches. Thirdly, the participants were both postgraduate and undergraduate students from a variety of academic courses. All of these reasons make it hard to generalize about the effect of aptitude tests and standardized or international exams on the academic accomplishment.

### **2.2.2 Locally Developed English Exams Conducted to Test the Relationship between Academic Achievement and Language Performance**

The framework embraced by the European Association of Language Testing and Assessment suggests that the teachers of English language and examiners must deliver a proof for the validity of the evaluation instrument or the tests they create (EALTA, 2006). Such proof essentially contains predictive validity research and a few researchers have investigated the locally established EAP performance exams' predictive power. Black (1991) examined the recorded scores in two different courses of language skills as the variable of the predictive power : EASL 140, which is about skills of speaking, contains listening, vocabulary, presentations, pronunciation, and comprehension, and EASL 143, which is about developing the skills of writing and contains library-searching techniques, problematic grammatical structures, and writing term papers. Lee (2005) researched an EAP exam, which is designated as valuation of the performance that involves students to summarize and incorporate articles' content as well as the lessons' content. The exam also encourages a procedure of writing and comprises pre-writing, using a word processor, re-drafting, and receiving feedback. Predictive validity researches formed a variety of correlation coefficients. Cope (2011) records statistically weighty connections between the GPAs and the scores of two task-based programs of  $r=0.361$  and  $r=0.418$ , and single correlation which did not attain significance ( $r=0.156$ ). He understands the statistics as displaying weak to considerable associations, where correlations of  $r=0.40 - 0.70$  were considered significant.

Black (1991) discovered slight to moderate correlations between the students' GPA and the skills scores. The researcher noted that the correlation with the strongest value ( $r=0.392$ ,  $p<0.05$ ) for single group only elucidates 15% of the inconsistency, which indicates that the rest, or 85% may be the product of further aspects, for example, adaptableness, enthusiasm, organizational skills. In Lees' (2005) paper, the correlation coefficients between the first semester GPA and the placement performance test varied. It was negative for Technology ( $r= 0.213$ ) and Life Sciences ( $r= 0.548$ ), but positive for Humanities ( $r=0.350$ ) and Business ( $r=0.275$ ). Robison & Ross (1996) inspected the relationship between performance on academic research tasks and the English language proficiency. They reviewed the degree to which an Exam of indirect Library Skills Research and a placement exam of English language predicted success on an exam of direct Library Skills Research that characterizes a genuine task for undergraduates in university. The outcomes indicated that only the exam of English language has the predictive power of the EAP research skills of the undergraduates. Nonetheless, success on direct exam of Library Skills Research might be predicted in a better way by an amalgamation of a language proficiency exam as well as an indirect Library Research Skills exam.

Alternatively, Cho's (2003) exploration of the workshop based essay predictive validity, which intended to imitate as narrowly as potential the features of the real-life of the academic writing, displayed that the predictive influence of the based essay was almost the same as that of the timed essay once associated with the students' faculty assessment. As Cho (2003) records, depending on the workshop exam for placement aims leads to extra incorrect positive results, i.e. the students were believed to have aptitude when they do not have, while the timed essay scores

caused additional incorrect negative results, i.e. students were believed not to have aptitude when they do. As nothing of these outcomes is acceptable, the writer attentions were in contrast to conclude interpretations around test-takers' aptitudes grounded only on a writing sample. As for predictive validity researches of large-scale of EAP exams, it is problematic to generalize from such a diversity of valuation measurements, each describing EAP proficiency exam in a dissimilar way, using participants from various academic fields, diverse criterion measures, and including diverse sample sizes.

Furthermore, other influences in addition to language proficiency are likely to affect the academic accomplishment. Academic background, current support and teaching, and personal background (Feast, 2002), quantitative skills, learning strategies, motivation (Bridgeman & Cho, 2012), acculturation in addition to intelligence (Cope, 2011) may affect the achievement of the students at university level.

Prapphal (1990) directed a research to investigate the sub-tests of the Exam of National English Entrance predictive validity of academic success in Thailand in English for Academic Purposes and Freshmen General English courses at universities in two different cities. The research covered 264 students of science who had engaged in the National English Entrance Exam in the year 1982. The outcomes showed that the sub-tests altogether correlated meaningfully with the achievement of English language in university. Subsequently, because the content of the tests engaged English in general. The sub-tests account for additional adjustment with the courses of General English than with the Academic English courses. Prapphal (1990) proposed that exam design could take a part in expecting the academic accomplishment in English language in the future. Prapphal (1990) explored the

linking between the Exam of English for the Academic Purposes, which required an extra course specific items and extra content and the exam of General English, which meant to measure the students' capability in understanding general English. The designs of both tests were alike. The research involved 320 students at Chulalongkorn University. Important indirect connections between the English for Academic Purposes and subskills of the General English were observed. This study exhibited that the subskills of the English language are connected together, whatever the format and the layout of the exam are. A transmission of language skills between the contents (General English) and (English for Academic Purposes) is probable.

In another research by Prapphal (1990) with 100 students. The research was conducted to discover to what degree the English for the Academic Purposes subtests, the English for the Academic Purposes Test and the English Entrance AB tests, expected to predict the performance in university, which is characterized by GPAs. The findings disclosed that regardless of whether every one of the tests could anticipate the academic achievement, the English for the Academic Purposes tests were progressively effective and more successful when contrasted with the test of Overall English language.

Aa investigation conducted by Fenton, Strough, and Stofflet (2001) studied the predictive power of the Graduation Qualifying exam at Alaska State High School and Benchmark Exam on the Achievement Tests performance in California. The conclusions suggested a strong connection between performances on the Benchmark Exam or High School Qualifying exam Writing marks and Reading marks and accomplishment on the Achievement Tests Total Language and Arts marks and Total Reading marks, correspondingly.

Researchers of Educational Testing Unit, McCauley-Jenkins and Ramist Lewis (2002) carried out a research, which examined the correlations between students GPA and SAT II Subject Tests. The outcomes showed that English ensured an association of .51 with students GPA. This was the maximum correlation amongst the Tests Subjects in SAT II. Further, an investigation accompanied by Breland, Kubota and Bonner (1999) considered the connection among the marks on the Writing Test in SAT II and the success in writing of the students in the university first year. 222 participants with every necessary writing examples were involved; nevertheless, extra conditions were obtainable for some variables once associated with others. The conclusions of the research presented high associations between university course marks and Verbal score in SAT I. In addition, a great correlation was noticed in the Writing Test in SAT II. Nevertheless, the score of Essay Writing Test in the SAT had a lesser correlation for predicting course marks once paralleled with SAT II: Writing Test and SAT I Verbal score.

A work by Ayers and Heard (1988) measured the American College Test (ACT) validity in expecting the performance on the test of Pre-Professional Skills (PPST). PPST is an exam made to test writing proficiency as well as reading. ACT, which is a test for admission, involves science, social science, math, English sub-tests adding to a compound test score. 202 participant involved in this research. The students had taken the PPST exam. The students had finished the ACT exam as well. The results revealed that the score of ACT composite was the finest indicator in prediction of the success and performance on the tests of PPST. Sub-test scores, the scores of ACT composite, along with GPA in the courses of English enriched the prediction of success and accomplishment. The outcomes indicated that the ACT exam scores

have prediction power and they are a balanced predictor of achievement on the PPST.

An examination was directed by Alavi (2012) to discover the prescient legitimacy of conclusive English exams as a measure of accomplishment in national college selection test. The examination included an example of 42 students at pre-university level in various fields of study. The outcomes demonstrated that there was a positive connection between each of the exams and national college entrance English exam, independently and in mix. The whole speculation that the study brought were affirmed up in various level of noteworthiness.

It appears that most of the predictive validity studies measured the relationship between the English exams and the students' GPAs as an indicator of their English performance. As for standardized English exams such as IELTTS and TOEFL, they were correlated with the students' GPAs. In addition, the locally developed English exams were correlated with students' GPAs. Most of the studies were based on the GPAs as a measure of student performance. Most studies also did not pay attention to the measurement of the predictive validity of the achievements exams in the English language courses offered in the university. In addition, it seems that there is no research on the correlation between the scores of two achievement exams.

Subsequently, the present study seeks to fill in the gap in the literature about the connection between midterm and final exams in English courses, as well as to search the levels of predictive validity in these exams. Further, it also examines if the exam results of the improvement courses could predict the future performance of the students.

## **Chapter 3**

### **METHODOLOGY**

This chapter shows the methodology that was used to conduct this research. It includes six major parts: the research design, research questions, data collection instruments, data collection procedures, the method of data analysis, and data analysis procedures.

#### **3.1 Research Design**

A quantitative methodology was utilized in this research. Leedy and Ormrod (2001) noted that researchers seek clarifications and predictions that will generate to other persons and places when they conduct a quantitative research. The intent is to create, confirm, or validate relationships and to develop generalizations that contribute to a theory.

This study focused on finding the relationship between many sets of student's scores in the English improvement courses such as, midterm and final exams' scores, and the total grade of these courses. As well as the relationship between the final grades of the first set of the courses with the scores of exams and finale grades in the second set of the English improvement courses presented at English Language Teaching Undergraduate Program in Eastern Mediterranean University. The researcher used multiple methods of statistical analysis such as correlation coefficient to find the predictive power and validity of the exams scores and the relationships between the scores.



### **3.2 Research Context**

The study was conducted at the Department of Foreign Language Education (FLE) of the Faculty of Education at Eastern Mediterranean University (EMU) in North Cyprus.

The Department offers one (BA) undergraduate and two (MA and Ph.D.) graduate study programs; the first one is an undergraduate program (BA) prompting the Bachelor certificate of Arts in ELT. As stated by the ELT program curriculum for the BA students, the EFL Department offers courses that are considered effective for teaching performance and professional development such as classroom management, teaching language skills, approaches to ELT, linguistic foundation, research methods, and testing and evaluation ([www.fedu.emu.edu.tr](http://www.fedu.emu.edu.tr)).

This study focus on the English improvement courses which are offered in the first year of the undergraduate program (BA). The courses are (Contextual Grammar – I and Contextual Grammar - II, Advanced Reading and Writing - I and Advanced Reading and Writing - II, Listening and Pronunciation - I and Listening and Pronunciation - II, Oral Communication Skills - I and Oral Communication Skills - II, Vocabulary courses). The program offer half of them in the first (fall) semester and the other half in the second (spring) semester.

In addition, the Department of Foreign Language Education has decently preserved teaching excellence and research international standards in many different levels in all of the university programs. In 2014, the FLE Department BA program received official accreditation from the AQAS (Agency for Quality Assurance of the accreditation educational-programs) that is registered with the European quality

program for higher education. The mission of the Department of Foreign Language Education is to provide contemporary tertiary education, to enhance the efforts of innovations and professional developments in the academic research studies, as well as to train capable, confident, skillful, knowledgeable and creative experts who are anticipated to take a larger educational parts in the current modern world ([www.fedu.emu.edu.tr](http://www.fedu.emu.edu.tr)).

General information: The qualification goal of the BA Program is to train all-rounded modern language teachers of English. The central idea of the program is to offer a rigorous and comprehensive training for BA students so that they can obtain an adequate competence in English language teaching.

### **3.3 Participants**

The participants were 1716 students of the Department of Foreign Language Education at Eastern Mediterranean University.

### **3.4 Research Questions**

This study seeks to provide answers for these questions:

1. Do the exams in the improvement courses have predictive validity?
2. Is there a connection between the mid-term exams' scores and the final exams' scores in the English improvement courses?
3. Is there any relationship between the final exams' scores and the courses' total grades?

### **3.5 Data Collection Instruments**

Data was collected from the Department of Foreign Language archive. The scores of the mid-terms and finals exams of the students, as well as their total grades in the English improvement courses: (Contextual Grammar – I, Contextual Grammar - II, Advanced Reading and Writing - I, Advanced Reading and Writing - II, Listening and Pronunciation – I, Listening and Pronunciation - II, Oral Communication Skills – I, Oral Communication Skills - II, and Vocabulary courses) for the academic years 2013-2018, were computed.

### **3.6 Data Collection Procedures**

Having obtained the form of consent of the Ethical Committee of the Eastern Mediterranean University (Appendix A) and the approval of the head of the Department of Foreign Language Education (Appendix B), the data was collected throughout the Spring Semester of the Academic Year 2017-2018 from the archive in the Department of Foreign Language Education.

### **3.7 Data Analysis Instruments**

Statistical analyses was used with the student scores. Data was analyzed using quantitative statistics methods and presented using tables. Data analysis is the procedure of systematically relating statistical and/or logical techniques to designate and illustrate, condense and recap, and value data. Shamo and Resnik, (2003), stated that numerous analytical procedures provide a way of illustrating inductive explanations from data. The data was analyzed using SPSS (Statistical Package for Social Sciences), Microsoft Excel, and other statistical formulas. The test scores for the sampled English improvement courses and the total scores, for the selected courses for the study, were recorded. The two sets of examination scores were compared by calculating Pearson Correlation Coefficient, ( $r$ ). The aim of the analysis

was to reveal the level of relationship between the variables to clarify the positive, negative, or zero correlation between the exams scores. Qualitative data analysis (content analysis) was used to interpret the results in relation to the objectives of the study.

Moreover, other measures were calculated in this research for descriptive statistical analysis, means, standard deviation, skewness value, Kurtosis value, maximum, and minimum scores. For measuring the normality of the data distribution, Anderson-darling test was applied. As for linearity of the data, scatter plots were graphed and examined.

### **3.8 Data Analysis Procedures**

After removing the students with empty scores or zeros and removing some data outliers. The scores of the mid-terms and the finals exams, and the total grades of 1716 students in the English improvement courses, for the academic years 2013-18 were ordered and recorded. The data was three sets of scores, the scores of mid-term exam, the scores final exam, and the total grade of the 9 English improvement courses for five academic years 2013-14, 2014-15, 2015-16, 2016-17 and 2017-18. After the sample was selected and filtered, descriptive statistics was calculated for each data set. For the combined data skewness and Kurtosis values, were calculated to test the symmetry and the pattern of distribution of the data.

As for the Pearson Correlation coefficient( $r$ ) in the first stage, for each course the correlation was calculated between mid-term exam scores and final exam scores, then between final exam scores and the total grade of the courses for each course in each academic year individually in the five years. In the second stage, all the scores

for each course were combined together from all the five academic years. Then the correlation was calculated for all the scores in each course between mid-term exam scores and final exam scores, then between final exam scores and the total grade of the courses among the nine courses in all the five academic years.

As for all the correlation apprehensive with the questions of the research, the null hypothesis (H<sub>0</sub>: there is no correlation amongst the variables; mid-term and final scores and between final scores and total grade), was rejected for significant correlation at  $p < 0.05$  and  $p < 0.01$ . So, the alternative hypothesis (H<sub>1</sub>: there is a significant correlation between all the variables) was accepted. (The results are significant at  $p < 0.05$  and  $p < 0.01$ ).

## **Chapter 4**

### **RESULTS**

The chapter presents the outcomes of this study on the descriptive statistics of the data of the students' exams scores for each academic year, the findings of the correlations between the students' exam scores for each academic year and descriptive statistical analysis for the students' exams scores of each course in the five academic years all together. In the last section, the correlations between the students' exam scores for all the five academic years are showed.

#### **4.1 Descriptive Statistical Analysis for the Students' Exams Scores of Each Course for Each Academic Year**

In this sector, the outcomes of the analysis of the data (scores of the mid-terms, finals, and courses' total grades) are presented.

Tables 1, 2, 3, 4, and 5 below present the results of the descriptive statistical analysis for the exams scores of the nine English improvement courses in each academic year.

The means and the standard deviations for each exam set of scores in each academic year were presented in the tables.

Table 1: Descriptive Statistics for English Improvement Courses' Exams Scores for the Academic Year 2013-14

Courses	N	Exams	Mean of the scores	St. Deviation of the scores
ELTE 101	32	Mid-term	20.44	5.17
		Final	25.32	5.19
		Total Grade	59.7	14.06
ELTE 102	43	Mid-term	18.44	6.83
		Final	23.41	7.1
		Total Grade	56.8	18.02
ELTE 103	25	Mid-term	23.3	4.95
		Final	22.26	7.39
		Total Grade	60.8	16.14
ELTE 104	37	Mid-term	20.3	5.45
		Final	23.05	9.03
		Total Grade	57.74	19.5
ELTE 105	33	Mid-term	16.8	3.56
		Final	15.7	5.38
		Total Grade	57.12	16.81
ELTE 106	17	Mid-term	18.6	4.5
		Final	18.23	5.8
		Total Grade	60.8	20.02
ELTE 107	51	Mid-term	16.84	4.14
		Final	22.71	3.77
		Total Grade	64.48	15
ELTE 108	57	Mid-term	12.35	3.8
		Final	23.41	4.21
		Total Grade	65.1	13.14
ELTE 112	23	Mid-term	14.96	6.41
		Final	16.41	11.2
		Total Grade	56.2	24.4

Table 2: Descriptive Statistics for English Improvement Courses' Exams Scores for the Academic Year 2014-15

Courses	N	Exams	Mean	St. Deviation
ELTE 101	36	Mid-term	21.32	5.6
		Final	20.15	7.1
		Total Grade	57.13	15.66
ELTE 102	43	Mid-term	19.26	8.44
		Final	19.48	9.013
		Total Grade	54.06	21.6
ELTE 103	35	Mid-term	19.74	5.44
		Final	24.34	6.3
		Total Grade	64.52	15.85
ELTE 104	39	Mid-term	21.35	6.32
		Final	22.74	10.06
		Total Grade	60.54	19.11
ELTE 105	24	Mid-term	16.75	7.2
		Final	20.21	3.05
		Total Grade	67.9	15.15
ELTE 106	37	Mid-term	18.61	4.77
		Final	15.7	5.89
		Total Grade	64.1	14.93
ELTE 107	36	Mid-term	14.31	5.24
		Final	22.94	5.34
		Total Grade	65.83	16.5
ELTE 108	33	Mid-term	17.96	5.32
		Final	21.45	7.74
		Total Grade	67.47	20.52
ELTE 112	41	Mid-term	13.41	5.83
		Final	20.35	10.11
		Total Grade	59.59	22.13



Table 3: Descriptive Statistics for English Improvement Courses' Exams Scores for the Academic Year 2015-16

Courses	N	Exams	Mean	St. Deviation
ELTE 101	38	Mid-term	18.3	5.45
		Final	27.06	5.56
		Total Grade	58.97	10.94
ELTE 102	39	Mid-term	13.14	4.96
		Final	15.54	7.51
		Total Grade	59.18	21.55
ELTE 103	36	Mid-term	20.35	5.21
		Final	22.73	6.17
		Total Grade	61.93	13.14
ELTE 104	35	Mid-term	18.43	5.76
		Final	26.12	6.96
		Total Grade	69.01	16.82
ELTE 105	36	Mid-term	18.28	4.45
		Final	17.42	6.44
		Total Grade	63.21	15.01
ELTE 106	32	Mid-term	19.3	5.32
		Final	17.9	8.14
		Total Grade	61.72	21.45
ELTE 107	34	Mid-term	18.71	3.92
		Final	27.71	3.22
		Total Grade	81.97	11.22
ELTE 108	31	Mid-term	23.11	7.27
		Final	21.13	8.44
		Total Grade	76.15	24.13
ELTE 112	21	Mid-term	16.1	4.92
		Final	20.6	9.46
		Total Grade	57.74	20.21

Table 4: Descriptive Statistics for English Improvement Courses' Exams Scores for the Academic Year 2016-17

Courses	N	Exams	Mean	St. Deviation
ELTE 101	35	Mid-term	17.25	5.76
		Final	24.17	6.44
		Total Grade	69.16	19.57
ELTE 102	32	Mid-term	16.58	6.68
		Final	20.28	9.28
		Total Grade	64.93	24.58
ELTE 103	29	Mid-term	18.05	4.88
		Final	26.36	9.71
		Total Grade	62.31	18.8
ELTE 104	31	Mid-term	28.59	5.8
		Final	18.65	7.75
		Total Grade	73.03	21.41
ELTE 105	35	Mid-term	18.2	6.68
		Final	18.89	6.91
		Total Grade	63.34	23.24
ELTE 106	34	Mid-term	17.96	6.16
		Final	20.22	8.77
		Total Grade	60.6	20.91
ELTE 107	27	Mid-term	21.33	6.01
		Final	9.44	2.12
		Total Grade	77.28	19.82
ELTE 108	29	Mid-term	26.07	5.73
		Final	21.38	7.69
		Total Grade	64.45	19.11
ELTE 112	34	Mid-term	15.025	6.01
		Final	22.49	9.09
		Total Grade	61.97	22.73

Table 5: Descriptive statistics for English Improvement Courses' Exams Scores for the Academic Year 2017-18

Courses	N	Exams	Mean	St. Deviation
ELTE 101	46	Mid-term	16.2	5.055
		Final	23.61	6.66
		Total Grade	69.31	17.84
ELTE 102	60	Mid-term	18.73	7.3
		Final	22.88	6.90
		Total Grade	66.24	21.23
ELTE 103	46	Mid-term	21.22	4.1
		Final	22.11	9.35
		Total Grade	63.39	18.3
ELTE 104	58	Mid-term	17.45	4.85
		Final	16.22	6.33
		Total Grade	59.58	17.42
ELTE 105	50	Mid-term	20.61	5.1
		Final	19.66	8.05
		Total Grade	68.14	19.5
ELTE 106	63	Mid-term	21.5	17.13
		Final	21.09	5.66
		Total Grade	69.56	15.22
ELTE 107	41	Mid-term	15.93	3.65
		Final	23.55	3.6
		Total Grade	69.28	12.15
ELTE 108	58	Mid-term	16.18	5.84
		Final	22.38	5.12
		Total Grade	68.17	14.76
ELTE 112	64	Mid-term	16.42	5.28
		Final	25.41	8.3
		Total Grade	68.12	19.97

## 4.2 The Correlations between the Students' Exam Scores for Each Academic Year

In this section, the scores of the mid-terms exam in each course are correlated with the scores of the final exam, and then the scores of the final exams are correlated with the total grads of the course of the students in each academic year.

Tables 6, 7, 8, 9, and 10 below present the correlation coefficient( $r$ ) between English Improvement Courses' exams scores for each academic year.

Table 6: Correlation Coefficient ( $r$ ) between English Improvement Courses' Exams Scores for the Academic Year 2013-14

<b>Courses</b>	<b>Correlation between Mid-term and Final (<math>r</math>)</b>	<b>Correlation between Final and course's total grade (<math>r</math>)</b>	<b>Number (N)</b>
<b>ELTE 101</b>	0.618**	0.8967**	32
<b>ELTE 102</b>	0.7812**	0.9343**	43
<b>ELTE 103</b>	0.8044**	0.9617**	25
<b>ELTE 104</b>	0.7178**	0.9361**	37
<b>ELTE 105</b>	0.6581**	0.897**	33
<b>ELTE 106</b>	0.9351**	0.9029**	17
<b>ELTE 107</b>	0.615**	0.7982**	51
<b>ELTE 108</b>	0.1841	0.6698**	57
<b>ELTE 112</b>	0.7044**	0.9302**	23

\*\* The result is significant at  $p < 0.05$  and  $p < 0.01$

Table 6 shows that the highest correlation between a mid-term and final exam is  $r=0.9351$ , a strong positive correlation, which means that high score in med-term exam goes with high final exam score (and vice versa). This indicates that the mid-term exam has a strong predictive validity. As for the correlation between final and courses' total grade the highest correlation is  $r=0.9617$ , a strong positive correlation,

which means that the final exam for the course ELTE 103 had a high predictive power in this academic year. On the other hand, the lowest correlation is  $r = 0.1841$  which is although technically a positive correlation, the relationship between the scores is weak. The results for almost all variables is significant at  $p < 0.05$  and  $p < 0.01$ .

Table 7: Correlation Coefficient (r) between English Improvement Courses' Exams Scores for the Academic Year 2014-15

<b>Courses</b>	<b>Correlation between Mid-term and Final (r)</b>	<b>Correlation between Final and course's total grade (r)</b>	<b>Number (N)</b>
<b>ELTE 101</b>	0.751**	0.9293**	36
<b>ELTE 102</b>	0.7984**	0.8775**	43
<b>ELTE 103</b>	0.8368**	0.9591**	35
<b>ELTE 104</b>	0.6053**	0.9547**	39
<b>ELTE 105</b>	0.4574**	0.6817**	24
<b>ELTE 106</b>	0.5007**	0.8554**	37
<b>ELTE 107</b>	0.5503**	0.8406**	36
<b>ELTE 108</b>	0.7347**	0.8755**	33
<b>ELTE 112</b>	0.6389**	0.8919**	41

\*\* The result is significant at  $p < 0.05$  and  $p < 0.01$

As for table 7 the results for all variables is significant at  $p < 0.05$  and  $p < 0.01$ . All the scores have strong positive correlation, which indicates high predictive validity for the final and mid-term exams.

Table 8: Correlation Coefficient (r) between English Improvement Courses' Exams Scores for the Academic Year 2015-16

<b>Courses</b>	<b>Correlation between Mid-term and Final (r)</b>	<b>Correlation between Final and course's total grade (r)</b>	<b>Number (N)</b>
<b>ELTE 101</b>	0.2434	0.6738**	38
<b>ELTE 102</b>	0.6958**	0.9273**	39
<b>ELTE 103</b>	0.2963	0.8142**	36
<b>ELTE 104</b>	0.5777**	0.9041**	35
<b>ELTE 105</b>	0.5108**	0.8911**	36
<b>ELTE 106</b>	0.6176**	0.9151**	32
<b>ELTE 107</b>	0.319	0.7438**	34
<b>ELTE 108</b>	0.6754**	0.9391**	31
<b>ELTE 112</b>	0.5736**	0.9281**	21

\*\* The result is significant at  $p < 0.05$  and  $p < 0.01$

Table 8 shows that ELTE 101 and ELTE 107 mid-term exams have low predictive power; the correlation between mid-term and final exam is weak and not significant at  $p < 0.05$  and  $p < 0.01$ .

Table 9: Correlation Coefficient (r) between English Improvement Courses' Exams Scores for the Academic Year 2016-17

<b>Courses</b>	<b>Correlation between Mid-term and Final (r)</b>	<b>Correlation between Final and course's total grade (r)</b>	<b>Number (N)</b>
<b>ELTE 101</b>	0.6996**	0.8833**	35
<b>ELTE 102</b>	0.8245**	0.9473**	32
<b>ELTE 103</b>	0.7985**	0.9784**	29
<b>ELTE 104</b>	0.7696**	0.9658**	31
<b>ELTE 105</b>	0.8998**	0.9591**	35
<b>ELTE 106</b>	0.7876**	0.952**	34
<b>ELTE 107</b>	0.5138**	0.81**	27

<b>ELTE 108</b>	0.6812**	0.9215**	29
<b>ELTE 112</b>	0.8993**	0.9549**	34

\*\* The result is significant at  $p < 0.05$  and  $p < 0.01$

As for the academic year 2016-17, table 9 shows that all correlations are positive and strong as well as significant at  $p < 0.05$  and  $p < 0.01$ .

Table 10: Correlation Coefficient (r) between English Improvement Courses' Exams Scores for the Academic year 2017-18

<b>Courses</b>	<b>Correlation between Mid-term and Final (r)</b>	<b>Correlation between Final and course's total grade (r)</b>	<b>Number (N)</b>
<b>ELTE 101</b>	0.6283**	0.8791**	46
<b>ELTE 102</b>	0.8927**	0.9255**	60
<b>ELTE 103</b>	0.7675**	0.8937**	46
<b>ELTE 104</b>	0.7037**	0.888**	58
<b>ELTE 105</b>	0.6975**	0.9529**	50
<b>ELTE 106</b>	0.4109**	0.8444**	63
<b>ELTE 107</b>	0.589**	0.8232**	41
<b>ELTE 108</b>	0.4367**	0.7789**	58
<b>ELTE 112</b>	0.6776**	0.919**	64

\*\* The result is significant at  $p < 0.05$  and  $p < 0.01$

All the correlations in table 10 are positive and strong, and the results is significant at  $p < 0.05$  and  $p < 0.01$ . The Mid-term and final exams are valid in terms of prediction for the academic year 2017-18.

### **4.3 Descriptive Statistical Analysis for the Students' Exams Scores of Each Course in the Five Academic Years Together**

In this sector, the outcomes of the analysis of the data (scores of the mid-terms, finals, and courses' total grades) are presented.

Table 11 below shows the results of the descriptive statistical analysis for the exams scores of the nine English improvement courses for the students in the five academic years combined. The means, the standard deviations, the minimum score, the maximum score, the skewness, the Kurtosis, and AD-test for each exam set of scores for each course in all the five academic years.



Table 11: Descriptive Statistics for English Improvement Courses' Exams Scores for the Academic Years 2013-2018

Courses	N	Exams	Mean	St. Deviation	Minimum Score	Maximum Score	Skewness	Kurtosis	The distribution is normal with a confidence of (AD-test)
ELTE 101	187	Mid-term	18.53	5.67	5.5	33.5	0.123	-0.183	75.11%
		Final	24.04	6.57	0	36	-1.14	2.24	0%
		Total Grade	63.2	16.65	6.5	99.25	-0.307	0.139	45.04%
ELTE 102	217	Mid-term	17.45	7.26	2	34.5	0.00331	-0.798	1.88%
		Final	20.6	8.3	0	40	-0.5942	0.2716	0.04%
		Total Grade	60.48	21.63	4.5	98.5	-0.484	-0.3908	0.03%
ELTE 103	171	Mid-term	20.5	5.05	6	32	-0.1916	-0.2076	39.6%
		Final	23.44	8.01	0	42	-0.5469	0.8884	0.52%
		Total Grade	62.75	16.4	10	96.5	-0.3106	0.3489	2.29%
ELTE 104	200	Mid-term	20.6	6.6	0	35	0.2895	-0.3639	0.09%
		Final	20.86	8.72	0	37.5	-0.4972	0.5202	0.47%
		Total Grade	63.16	19.3	12	98.7	-0.4528	-0.2158	0.26%
ELTE 105	178	Mid-term	18.43	5.5	0	29.5	-0.6612	0.9889	0.63%
		Final	18.38	6.6	0	30	-0.6944	0.7795	0.1%
		Total Grade	64.12	18.65	8	100	-0.6563	0.2358	0.05%
ELTE 106	183	Mid-term	19.6	5.07	3.8	29.5	-0.4256	-0.2474	1.63%
		Final	19.01	7.07	0	34	-0.7851	0.8581	0%
		Total Grade	64.6	18.12	9.8	97	-0.6385	0.1202	0.04%
ELTE 107	189	Mid-term	17.13	4.99	0	29.5	-0.4839	0.153	0.2%
		Final	21.93	6.58	0	30	-0.9459	0.4023	0%
		Total Grade	70.8	16.17	7	98.5	-0.7085	1.126	0.06%
ELTE	208	Mid-term	17.8	7.22	0	35	0.0159	0.0204	0.86%

108		term							
		Final	22.2	6.34	0	35	-1.748	4.132	0%
		Total	67.89	17.83	0	97	-0.8706	1.5	0.07%
		Grade							
ELTE	183	Mid-	15.3	5.69	2	25	-0.0979	-1.044	0.03%
112		term							
		Final	22.1	9.73	0	35	-0.8246	-0.3151	0%
		Total	62.37	21.77	6	99.8	-0.4701	-0.5236	0.06%
		Grade							

Table 11 summarizes a statistical description of the data for each course separately during the five academic years. The mean, the standard deviation, minimum and maximum scores were recorded in each course for the mid-term exam, the final exam and the total grades. Skewness and Kurtosis values were calculated to give an idea about the shape and the symmetry of the distribution of each set of the data.

The average value of the skewness is -0.52, which indicates that the data are symmetrical for almost all the datasets as it is shown in the Figures 1 and 2. If the skewness value is between 0.5 and -0.5 so it has symmetrical distribution. The average of the kurtosis value for all the datasets is 0.38 a positive value indicates that datasets have more weight in the tails. If a dataset has a positive kurtosis, it has more in the tails than the normal distribution. If a dataset has a negative kurtosis, it has less in the tails than the normal distribution. Most often, kurtosis is measured against the normal distribution. If the kurtosis is close to 0, then a normal distribution is often assumed.

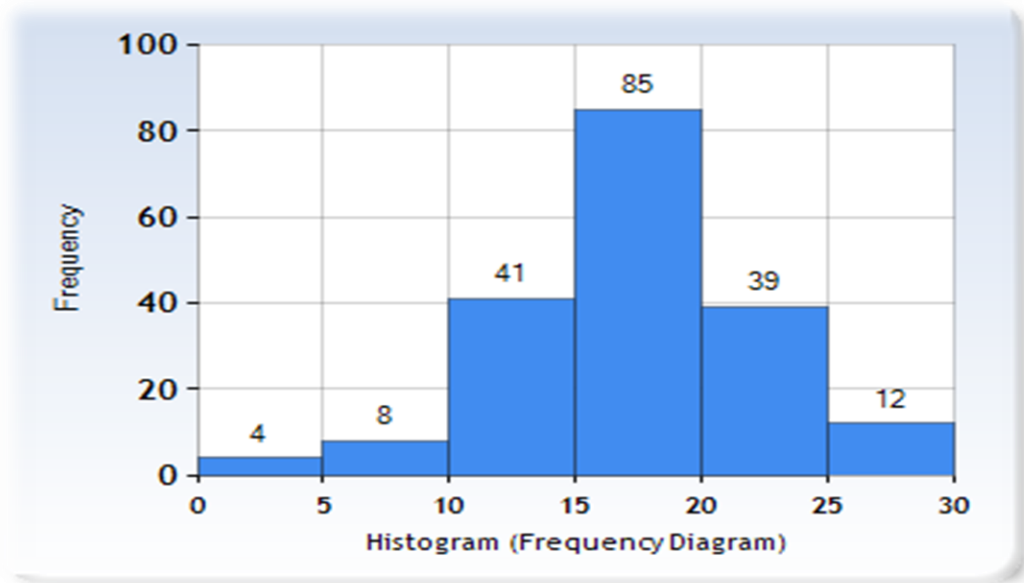


Figure 1: ELTE 107 mid-terms' exam scores distribution

ELTE 107 mid-term exam scores distribution with skewness -0.48 and kurtosis 1.6. The left-hand tail is longer than the right-hand tail. The data are symmetrical.

Most of the datasets distribution is symmetrical and have more weight in the tails as is showed in the Figure above with average skewness value -0.52 and average Kurtosis value 0.38 as both values close to zero.

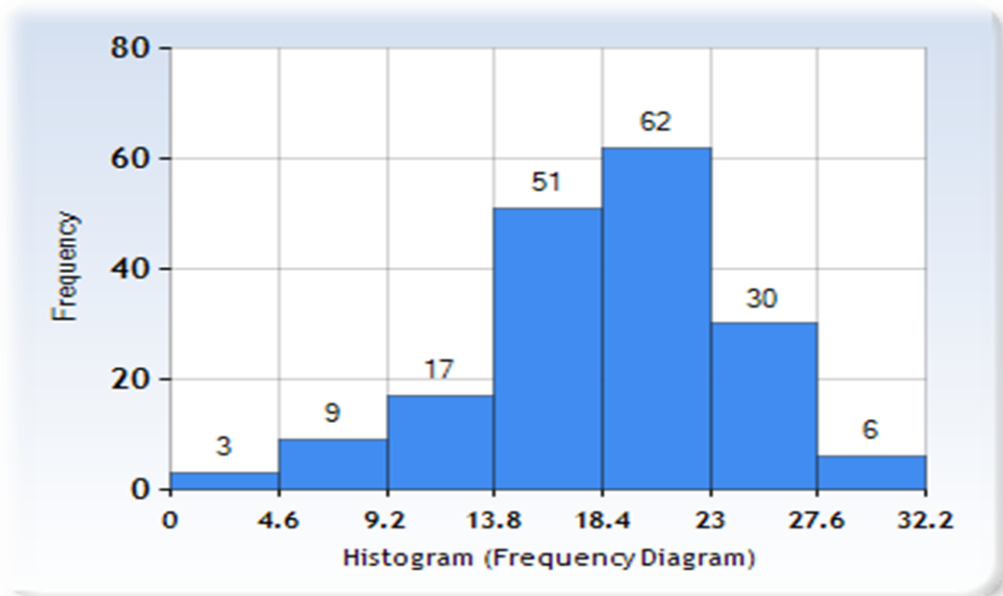


Figure: 2 ELTE 105 mid-terms' exam scores distribution

The skewness of the data is  $-0.6612$ . The kurtosis is  $0.9889$ . Both values are close to zero, as one would expect for a normal distribution. These two numbers represent the "true" value for the skewness and kurtosis since they were calculated from all the data.

Cumulative distribution and empirical distribution are given (Appendix D) which can be representative for all the datasets.

#### 4.4 The Correlations between the Students' Exam Scores for all the Five Academic Years

In this section, the scores of the mid-terms exam in each course were correlated with the scores of the final exam, and then the scores of the final exams were correlated with the total grads of the course of the students in all the academic years combined.

Table 12 below present the correlation coefficient(r) between English Improvement Courses' exams scores for the academic year 2013-18.

Table 12: Correlation Coefficient (r) in English Improvement Courses' Exams Scores for the Academic Year 2013-18

<b>Courses</b>	<b>Correlation coefficient between Mid-term score and Final score (r)</b>	<b>Correlation coefficient between Final score and course's total grade (r)</b>	<b>Number (N)</b>
<b>ELTE 101</b>	0.4823**	0.7698**	187
<b>ELTE 102</b>	0.8**	0.8636**	217
<b>ELTE 103</b>	0.5832**	0.9**	171
<b>ELTE 104</b>	0.4945**	0.821**	200
<b>ELTE 105</b>	0.6372**	0.91**	178
<b>ELTE 106</b>	0.6178**	0.8664**	183
<b>ELTE 107</b>	0.0527	0.3738**	189
<b>ELTE 108</b>	0.3351**	0.8097**	208
<b>ELTE 112</b>	0.6932**	0.9179**	183

\*\* The result is significant at  $p < 0.05$  and  $p < 0.01$

Table 12 presents the correlation between the mid-term exam scores and final exam scores, as well as between the final exam scores and the total grades for these courses for 1716 students distributed on nine courses in the academic years 2013-18. The results showed that the final exams in the English improvement courses have more predictive power than the mid-term exams in the five years.

## Chapter 5

### DISCUSSION AND CONCLUSION

This chapter discusses the results of this research based on the questions of the research. A discussion and summary of the findings of this research were presented, followed by the implications of the study, limitations, and suggestions for further research.

#### 5.1 Discussion of Results

In this section, the findings of the data analysis of this research were discussed in line with the research questions.

##### 5.1.1 Question 1: Do The Improvement Courses' Exams Have Predictive validity?

The first question relates to the predictive validity of the main two exams in the improvement courses (mid-term and final), thus the correlation between the scores was measured for five academic years. Based on the data analysis results the correlations between the mid-term and final exam was calculated to investigate the predictive validity of the mid-term exams for the improvement courses, as well as the correlations between the final exams and the total grades of the courses to examine the predictive validity of the final exams.

For the academic year 2013-14, the correlations between the mid-term exams and final exams were strong and positive for almost all the courses except for one non-significant weak relationship for the mid-term exam in ELTE 108 course  $r=0.1841$ .

The highest correlation for this academic year was seen in the mid-term exam in ELTE 106  $r=0.9351$ . The correlations average for the mid-term exams for this year was  $r= 0.595$ , which is a moderate positive correlation. The correlations between the final exams and the total grades of the courses was very strong and positive the average of the correlations was  $r=0.8807$  and all the correlations were significant. The highest correlation was  $r=0.9617$  which is for the final exam of ELTE 103 course. The weakest correlation was  $r=0.6698$  which was for the final exam of ELTE108 course. In this year (2013-14), the final exams have stronger predictive validity than the mid-term exams. The numbers of students in each course for this academic year were not big with average of  $N= 35.3$  student for each course with total number of  $N=318$  students in the all courses.

In the academic year 2014-2015, the correlations between the mid-term exams and final exams vary. The courses ELTE 101  $r= 0.751$ , ELTE102  $r= 0.7984$ , ELTE 103  $r= 0.8368$ , ELTE 108  $r= 0.7347$ , have strong positive correlation, but the courses ELTE 106  $r= 0.5007$ , ELTE 107  $r= 0.5503$ , ELTE 112  $r= 0.6389$ , have moderate and positive relationship, while ELTE 105 has a weak correlation  $r= 0.4574$ . The correlations average for the mid-term exams is  $r= 0.6017$  a moderate positive which indicate a moderate predictive validity for the mid-term exams in the courses for this year. The correlations between the final exams and the total grades of the courses were very strong and positive which means that the finals exam for the courses in this year have a strong predictive power, the average of the correlations for the final exams was  $r= 0.8739$ . All the correlation were significant and valid for the total number of students  $N=324$  and average student number for each course  $N=36$ .

The results obtained from the data analysis, revealed that the exams predictive power in academic year 2015-16 in general is lesser than the previous years. In this year, most of the correlations between the mid-term exams and final exams were weak and non-significant at  $p < 0.05$  and  $p < 0.01$  ( $r = 0.2434$ ,  $r = 0.2963$ ,  $r = 0.319$ ). The average mid-term exams correlation is  $r = 0.5010$  with moderate predictive value. Supplementary, the values of the correlation between the final exams and the total grades of the courses were very strong and positive, the average of the correlations was  $r = 0.8596$  which indicated high and strong predictive validity. The total number of students in this year was  $N = 302$ , and the average number of students for each course was  $N = 33.5$ .

As for the academic year 2016-17, according to the results presented in table 9, both mid-term exam and final exam have high strong positive correlations in all courses. The average correlated value for the mid-term exams with the final exams was  $r = 0.7637$ . While the average of the correlations for final exam correlated with the total grade was  $r = 0.7179$ . All the correlations values were significant and the total number of students was  $N = 286$  with average number of students for each course being  $N = 31.7$ . Thus, both mid-term exams and final exams have high predictive validity.

In the academic year 2017-18, the correlations between the mid-term exams and final exams were strong and positive for almost all the courses except for two weak relationship for the mid-term exam in ELTE 108 course  $r = 0.4367$  and ELTE 106  $r = 0.4109$ . The highest correlation for this academic year was for the mid-term exam in ELTE 102  $r = 0.8927$ . The average of the correlations for the mid-term exams in this year was  $r = 0.6448$  which is a moderate strong positive correlation. The correlations



between the final exams and the total grades of the courses were very strong and positive. The average of the correlations was  $r = 0.8783$  and all the correlations were significant. The highest correlation was  $r = 0.9529$  which is for the final exam of ELTE 105 course. The weakest correlation was  $r = 0.7789$  which is for the final exam of ELTE108 course. In this year, the final exams and the mid-term exams have strong predictive validity. The numbers of students in each course for this academic year were big compared to the previous years with average of  $N = 54$  students for each course with total number of  $N = 486$  students in the all courses.

The analysis of the correlation for the combined data for all the students in the five academic years (Table 12) was compared and correlated, regarding the courses not the years. The outcomes revealed that the strongest association between mid-term and final was  $r = 0.8$  in ELTE 102 courses, the weakest correlation was  $r = 0.0527$  in ELTE 107 courses. In addition, the average of the correlations for mid-term exams is  $r = 0.521$ , a moderate positive relationship which indicates a moderate predictive validity for the mid-term exams in the English improvement courses. As for the correlation between final exam and total grade, the strongest correlation was  $r = 0.9179$  in the ELTE 112 courses, the weakest correlation was  $r = 0.3738$  in the ELTE 107 courses. Moreover, the average of the correlations between final exam and total grades was  $r = 0.8035$ , a positive strong correlation which indicates a high predictive validity for the final exams in the English improvement courses. All the results were significant except  $r = 0.0527$  in the mid-term exam for the ELTE 107 courses. The findings showed that the final exams have more predictive validity than the mid-term exams. Furthermore, the exams of the course ELTE 107 have the weakest predictive validity among all the courses.

After reviewing the results of the data analysis for the five years individually and for the exams in the data combined, the results showed that the majority of the exams in the improvement courses have predictive validity.

The results showed that the exams in the English improvement courses (mid-terms and finals) have predictive power. The mid-terms scores were correlated with the final exams scores to discover if the mid-terms scores could predict the performance of the students in the final exams, and the final exams were scores correlated with the courses' total grades to discover if the final exams scores could predict the performance of the students in the final exams. After analyzing the data to find the correlations, the strong correlations between the exam scores revealed that the exams scores could predict the performance of the students in future exams and tests, and that the majority of these scores represent a real performance of the students.

#### **5.1.2 Question 2: Is there a Connection between the Mid-term Exams' Scores and the Final Exams' Scores?**

The analysis of the data proved that there is a positive relationship between the scores of mid-term exams and the final exams' scores. This means that as the score of a mid-term exam increased, the score of the final exam also increased.

#### **5.1.3 Question 3: Is there any Relationship between the Final Exams' Scores and the Courses' Total Grades?**

The analysis of the data proved that there is a positive relationship between the scores of the final exams and the courses' total grades. This mean that as the score of a final exam increased the scores of the courses' total grade also increased.

## **5.2 Conclusion**

The results of this study showed that the majority of the exams in the improvement courses have predictive validity. This was revealed through the strong correlations between the exam scores, which showed that the exams scores could predict the performance of the students in future exams and tests, and that the majority of these scores represent a real performance of the students. The findings revealed that there is positive relationship between the scores of mid-term exams and the final exams' scores and proved that there is a positive relationship as well between the scores of the final exams and the courses' total grades.

## **5.3 Implications of the Study**

This study will be useful in evaluating the exams of the improvement courses in the Department of Foreign Language in the Eastern Mediterranean University. The study will also contribute to the improvement of the exams in the program. It will contribute in finding the relationship between the results of the students in the midterms and final exams and their final grades in the courses. The results will benefit the evaluation of these exams. Further, this study will help to highlight the relationship between the exams scores used in English improvement courses, and the total grades of these courses, and to know if these exams have predictive validity as well as if they can predict the future performance of the students in future courses. To know whether students are on the right track or not. The instructors who offer the courses can reorganize or readapt their teaching to help those students who might not be on the right tract to improve their learning before they reach the end of the term. This study will give teachers and professors more reasons to take predictive validity into consideration and use it as a method to evaluate their achievement exams.

## **5.4 Limitations**

While conducting this study some limitations appeared. The first one relates to the sample-size of the study, as the amount of undergraduate students in ELT department was not as big as this study required. Also in some years, the number of students in some courses were less than 30. A bigger sample size would make the findings of the study more reliable and powerful.

The second limitation was the fact that the researcher was not knowledgeable about the nature of the exams and the questions. Further knowledge of the exam nature and questions would have helped more with the predictive validity.

Last, the study used only quantitative research method approach. The use of mixed method might have help to add more to the information gotten from the quantitative data analyzed.

## **5.5 Suggestions for Further Research**

Further correlation studies with larger sample are suggested. Further research should be carried out using a mixed-methods approach in order to identify the causes of the relationship between the exam scores, which will allow for more explanation for the findings. Asking teachers and analyzing the exam items is also suggested to make the results of the study more powerful.

## REFERENCES

- Al Hajr, F. (2014). The Predictive Validity of Language Assessment in a Pre-University Programme in two Colleges in Oman: A Mixed-Method Analysis. *International Multilingual Journal of Contemporary Research*, 2(2), 121-147
- Alavi, T. (2012). The Predictive Validity of Final English Exams as a Measure of Success in Iranian National University Entrance English Exam. *Journal of Language Teaching and Research*, 3(1), 224-228.
- Bachman, L. F, and Palmer. A. (2000). *Language Testing in Practice*. Oxford: Oxford University Press.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford university press.
- Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge: Cambridge University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests* (Vol. 1). Oxford University Press.
- Baker, D. (1989). *Language testing: A critical survey and practical guide*. London: Edward Arnold.

- Breland, H. M., Kubota, M. Y., & Bonner, M. W. (1999). The Performance Assessment Study in Writing: Analysis of the SAT® II: Writing Subject Test. *ETS Research Report Series, 1999(2)*, i-18.
- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment* (Vol. 17). Sage publications.
- Charney, D. (1984). The validity of using holistic scoring to evaluate writing: A critical overview. *Research in the Teaching of English, 65-81*.
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological assessment, 7(3)*, 309.
- Cope, N. (2011). Evaluating locally-developed language testing. *Australian Review of Applied Linguistics, 34(1)*, 40-59.
- Cotton, F., & Conrow, F. (1998). An investigation of the predictive validity of IELTS amongst a group of international students studying at the University of Tasmania. *International English Language Testing System (IELTS) Research Reports 1998: Volume 1*, 72.
- Crowder, M. J. (2017). *Statistical analysis of reliability data*. Routledge.
- Davies, A. (1990). *Principles of Language Testing*. Oxford: Blackwell.

- Dooley, P. (1999). An investigation into the predictive validity of the Battery Test as an indicator of future academic success. *Journal of New Generation*, 6(2), 60-96.
- Dooley P. & Oliver D. (2002). An investigation into the predictive validity of the IELTS Test as an indicator of future academic success. *Prospect*, 17(1) 36-54
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*: Oxford University Press Oxford.
- Elder, C. (1993). Language proficiency as a predictor of performance in teacher education. *Melbourne Papers in Language Testing*, 2(1), 72-95.
- European Association for Language Testing and Assessment (EALTA). (2006). Guidelines for good practice in language testing and assessment. Retrieved November 4, 2008, from <http://www.ealta.eu.org/guidelines.htm>
- Feast, V. (2002). The impact of IELTS scores on performance at University. *International Education Journal*, 3(4), 70-85.
- Fox, J. (2004). Test decisions over time: Tracking validity. *Language Testing*, 21(4), 437-465.
- Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The qualitative report*, 8(4), 597-606.

- Goodwin, L. D., & Leech, N. L. (2003). The meaning of validity in the new standards for educational and psychological testing: Implications for measurement courses. *Measurement and evaluation in Counseling and Development, 36*(3), 181-191.
- Graham, J. J. (1987). English language proficiency and the prediction of academic success. *TESOL Quarterly, 21*(3), 505-521.
- Graves, D. H. (2002). *Testing is Not Teaching: What Should Count in Education*. Heinemann, PO Box 6926, Portsmouth, NH 03802-6926.
- Harrison, A. (1983). *A language Testing Handbook*. London: Mcmillan.
- Hughes, A. (2003). *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- Hughes, A. (2003). *Testing English for Language Teachers*. Cambridge. UK: Cambridge University Press.
- Huong, T. (2001). The predictive validity of the international English language test system (IELTS). *The Post Graduate Journal of Education Research, 2*(1), 66-96.
- Huong, T. T. T. (2001). The predictive validity of the international English language testing system (IELTS) test. *Post-Script, 2*(1), 66-96.



- Jochems, W. (1991). Effects of learning and teaching in a foreign language. *European Journal of Engineering Education, 16(4)*, 309-316.
- Jochems, W., Sinppe, J., Jan Smid, H., & Verweij, A. (1996). The academic progress of foreign students: Study achievement and study behaviour. *Higher Education, 31(3)*, 325-340.
- Kane, M. (2011). *Validating score interpretations and uses: Messick lecture*. A paper presented in Language Testing Research Colloquium, (pp.3-17). UK: Cambridge.
- Kerstjens, M., & Nery, C. (2000). Predictive validity in the IELTS test: A study of the relationship between IELTS scores and students' subsequent academic performance. *International English Language Testing System (IELTS) Research Reports 2000: Volume 3*, 85.
- Maleki, A., & Zanjani, E. (2007). A survey on the relationship between English Language proficiency and the academic achievement of Iranian EFL students. *Asian EFL Journal, 9(1)*, 86-96.
- Marvin, L., & Simner, C. (1999). Postscript to the Canadian Psychological Association's position statement on the TOEFL. Retrieved Dec, 5, 2002.
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational researcher, 18(2)*, 5-11.

- Messick, S. (1995). Standards of Validity and the Validity of Standards in Performance Assessment. *Educational Measurement*, 14 (4), 5-8.
- Mohammadi, M. (2009). The Predictive Validity of Islamic Azad University's Entrance Examination: Does Access Mean Success?. *The Journal of Modern Thoughts in Education*, 4(4), 59-72.
- Mousavi, S. A. (1999). *A Dictionary of Language Testing*. Tehran, Iran: Rahnama publications.
- National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. National Academies Press.
- Prapphal, K. (1990). "The Relevance of Language Testing Research in the Planning of Language Programmes." Retrieved Jan 5, 2003.
- Ramist, L., Lewis, C., & McCauley-Jenkins, C. (2002). Validity of the SAT II Science Tests. *Science Insights*, 6(5).
- Rumsey, L. W. (2013). *Shedding light on the predictive validity of English proficiency tests in Predicting academic success*. PhD Thesis, Newcastle University
- Stofflet, F., Fenton, R., Strough, T. (2001). *Construct and Predictive Validity of the Alaska State High School Graduation Qualifying Examination: First*

*Administration*. A paper presented at the 2001 American Educational Research Association Convention.

Woodrow, L. (2006). Academic success of international postgraduate education students and the role of English proficiency. *University of Sydney papers in TESOL, 1*(1), 51-70.

Yen, D., & Kuzma, J. (2009). Higher IELTS score, higher academic performance? The validity of IELTS in predicting the academic performance of Chinese students. *Worcester Journal of Learning and Teaching, 3*(3).

## **APPENDICES**

# Appendix A: Approval Letter from the Ethical Committee (BAYEK)



**Eastern  
Mediterranean  
University**  
"For Your International Career"

P.K.: 99628 Gazimağusa, KUZEY KIBRIS /  
Famagusta, North Cyprus,  
via Mersin-10 TURKEY  
Tel: (+90) 392 630 1995  
Faks/Fax: (+90) 392 630 2919  
bayek@emu.edu.tr

Etik Kurulu / Ethics Committee

**Sayı:** ETK00-2018-0166  
**Konu:** Etik Kurulu'na Başvurunuz Hk.

31.05.2018

Yaseen Zeedan

Yabancı Diller Eğitimi Bölümü  
Yüksek Lisans Öğrencisi

Doğu Akdeniz Üniversitesi Bilimsel Araştırma ve Yayın Etiği Kurulu'nun 07.05.2018 tarih ve 2018/58-07 sayılı kararı doğrultusunda, *The Investigation of the Predictive Validity of ELT Improvement Courses' Exams in ELT Undergraduate Program* adlı çalışmanızı, Yrd. Doç. Dr. İlkay Gilanlıoğlu'nun danışmanlığında araştırmanız, Bilimsel ve Araştırma Etiği açısından uygun bulunmuştur.

Bilginize rica ederim.

**Doç. Dr. Şükrü TÜZMEN**  
Etik Kurulu Başkanı

ŞT/ba.

[www.emu.edu.tr](http://www.emu.edu.tr)

## **Appendix B: Permission Letter from the Department of Foreign Language Education**

Permission letter

Date: 30/4/2018

Dear Mr. Assist. Prof. Dr. Javanshir Shibliyev

I am Yaseen Zeedan and I am an MA student in ELT program at Eastern Mediterranean University undertaking a study on the topic "The Investigation of the Predictive Validity of ELT Improvement Courses' Exams in ELT Undergraduate Program." You are kindly requested to provide information about the midterm and final exams results and the courses grades as well as the students' GPAs and CGPAs in the English improvement courses: (Contextual Grammar - I, Contextual Grammar - II, Advanced Reading and Writing - I, Advanced Reading and Writing - II, Listening and Pronunciation - I, Listening and Pronunciation - II, Oral Communication Skills - I, Oral Communication Skills - II, and Vocabulary courses). Related to the academic years 2014- 2018. In addition, you are kindly requested to give the permission to conduct the study in the department of Foreign Language Education in Eastern Mediterranean University. All information will be treated strictly confidential and for academic purposes.

Sincerely,

Yaseen Zeedan

Master Student in the Department of Foreign Language Education

## Appendix C: Data Collection Permission Letters

To: Hüseyin Ünsal Yetiner

Registrar

From: Assist. Prof. Dr. İlkay Gilanlioğlu

FLE Department, Faculty of Education

08 June 2018



I am writing to kindly ask you to grant my MA student Yaseen Zeedan (Student number: 16500226) permission to use the following grades for the purposes of his thesis. He has already received written approval from Ethics Committee.

The midterm and final exam results, and the overall grades (out of 100), as well as the students' GPAs and CGPAs for the English language improvement courses (see the list below) related to the academic years 2013-2018.

ELTE 101 Contextual Grammar I

ELTE 102 Contextual Grammar II

ELTE 103 Advanced Reading and Writing I

ELTE 104 Advanced Reading and Writing II

ELTE 105 Listening and Pronunciation I

ELTE 106 Listening and Pronunciation II

ELE 107 Oral Communication Skills I

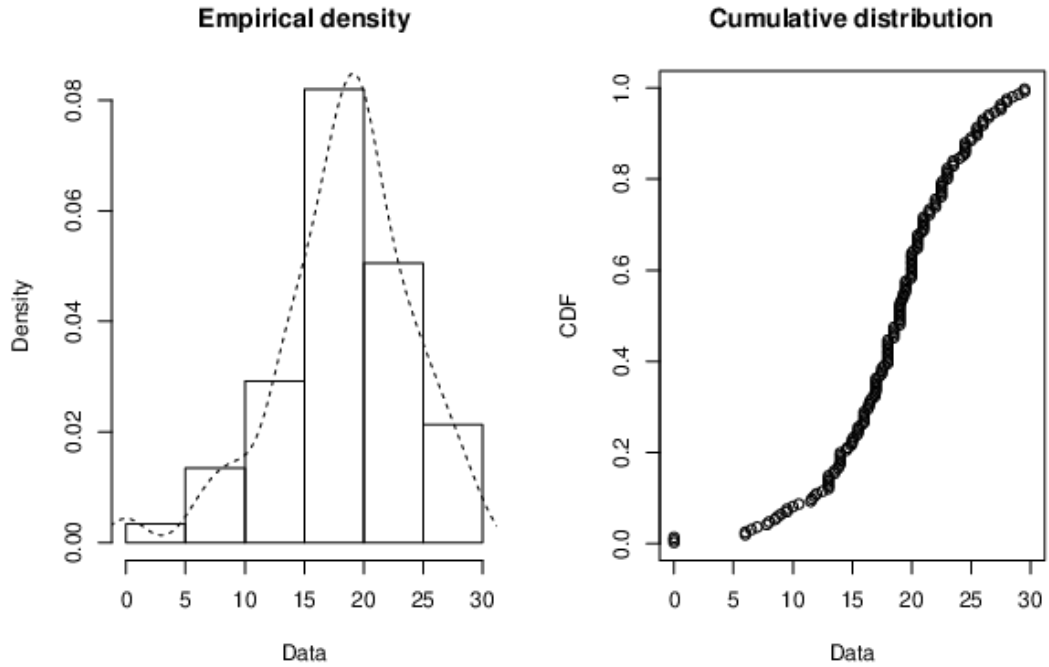
ELTE 108 Oral Communication Skills II

ELTE 112 Vocabulary

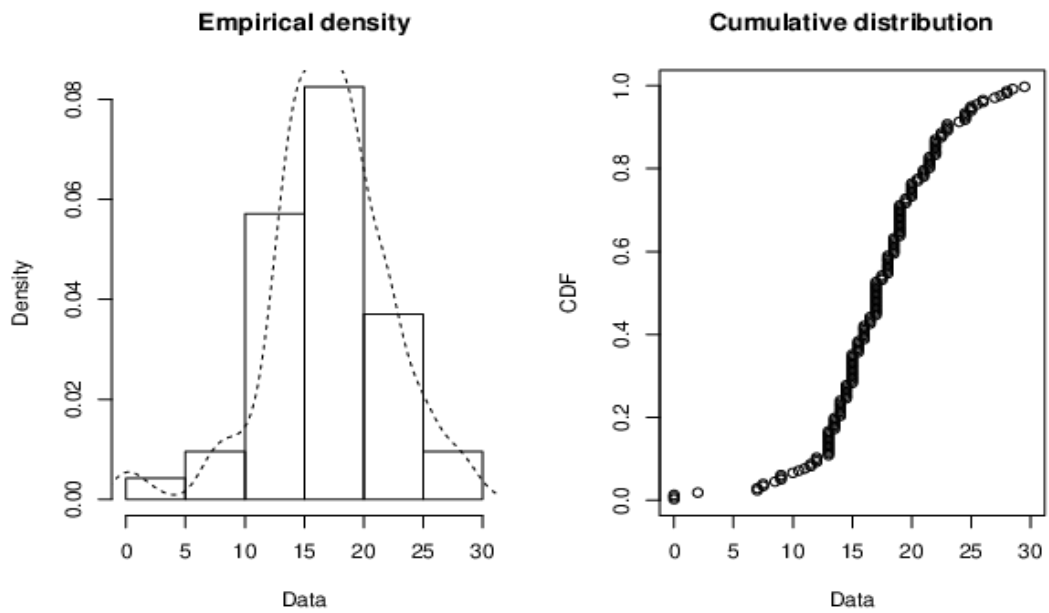
*Ilkay Gilanlioğlu*  
*J. Y. Yetiner*  
*S. W. Beyaz*  
08.06.2018

A circular blue ink stamp from Eastern Mediterranean University, Faculty of Education, FLE Department. The stamp contains the text "Eastern Mediterranean University", "Faculty Of Education", and "FLE DEPARTMENT". The date "08.06.2018" is written in the center of the stamp.

## Appendix D: Exams' Scores Distributions Histograms



ELTE 105 mid-terms' exam scores distribution



ELTE 107 mid-terms' exam scores distribution