

**THE IMPACT OF COMPUTER ASSISTED ASSESSMENT
ON THE EXAM SUCCESS AND ATTITUDES OF
PROSPECTIVE ENGLISH TEACHERS**

**BİLGİSAYAR TEMELLİ ÖLÇME-DEĞERLENDİRMENİN
İNGİLİZCE ÖĞRETMEN ADAYLARININ SINAV BAŞARISI
VE TUTUMU ÜZERİNE ETKİSİ**

Tuba ÖZTURAN

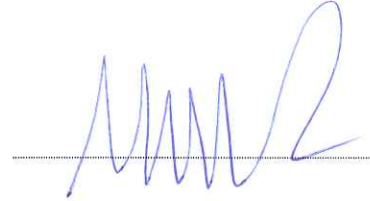
Submitted to the Graduate School of Educational Sciences of Hacettepe
University in Partial Fulfillment of the Requirements for the Master Degree in
English Language Teaching

2016

This is to certify that we have read this thesis, entitled “*The Impact of Computer Assisted Assessment on the Exam Success and Attitudes of Prospective English Teachers*”, and that in our opinion it is fully adequate, in scope and quality, as a thesis for the Degree of Master in the Program of English Language Teaching.

Examining Committee Members

Chairman Prof. Dr. Mehmet DEMİREZEN



Member
(Supervisor) Asst. Prof. Dr. Hüseyin ÖZ



Member Prof. Dr. İsmail Hakkı MİRİCİ



Member Assoc. Prof. Dr. İsmail Hakkı ERTEN



Member Assoc. Prof. Dr. Paşa Tevfik CEPHE



APPROVAL OF THE GRADUATE SCHOOL OF EDUCATIONAL SCIENCES

This is to certify that this thesis was approved by the aforementioned examining committee members on January 28, 2016 in accordance with the relevant articles of the Rules and Regulations of Hacettepe University Graduate Education, and accepted as a Master Thesis in the program of English Language Teaching by the Board of Directors of the Graduate School of Educational Sciences on// 2016.

Prof. Dr. Berrin AKMAN
Director

THE IMPACT OF COMPUTER ASSISTED ASSESSMENT ON THE EXAM SUCCESS AND ATTITUDES OF PROSPECTIVE ENGLISH TEACHERS

Tuba ÖZTURAN

ABSTRACT

Assessment is an indispensable part of education, and it is really crucial to prepare valid, reliable and practical tests, to provide a relaxed exam environment for test-takers and to grade them in an objective way. Today, with the introduction of computers into education, technology provides a number of opportunities in the assessment of learning and teaching outcomes in education, and one recent popular trend is computer-assisted assessment, which constitutes the base of the present study.

This study aimed to investigate the impact of computer assisted assessment on the perceptions and course achievement performance of prospective English language teachers. In this regard, mixed method research design was utilized. The quantitative data were gathered from a questionnaire with 37 items and an achievement test with 60 items. On the other hand the qualitative data were gotten from interview sessions.

97 student teachers who enrolled in the department of English Language Teaching at Hacettepe University were randomly assigned to experimental and control groups. In order to collect quantitative data, the experimental group took the exam on computer while the control group took it on paper, and then, all participants were wanted to rate a questionnaire in order to define their attitudes toward computer assisted assessment systems. The questionnaire, which was based on Technology Acceptance Model, had five dimensions in it such as intention, perceived usefulness, perceived ease of use, anxiety and computer attitude. In addition, students' demographic information like their gender, grade point average, daily use of computer and computer literacy were collected with an aim to define any probable relation between them. As for the qualitative data, 26 randomly chosen students were interviewed in order to support the statistical findings and reveal any other suggestions.

The results of the questionnaire showed that the experimental group had intention to computer assisted assessment while the control group showed no interest in the system. Likewise, the majority of the students in the experimental group thought that the system had increased their performance during the exam. Also, many students in that group assumed that they could learn how to navigate the system without any difficulty if computer assisted assessment was offered to them.

The students' attitudes toward computer assisted assessment may change up to their gender, grade point average, daily use of internet/ computer and computer literacy. The results displayed that gender was not statistically significant variable in terms of the students' attitudes toward computer assisted assessment. However, a moderate positive relation between computer literacy and perceived ease of use was found, which meant that more proficient computer users could use the system without challenges. Computer literacy, on the other hand, had a moderate negative correlation with anxiety, so proficient computer users showed lower anxiety comparing to the students who were non-proficient. In addition, grade point average and perceived ease of use had a weak negative relation, and lastly, there was a moderate positive relation between perceived ease of use and daily use of computer.

As for the groups' choices about test administration mode, the student teachers in the control group were strongly fond of taking paper-and-pencil based tests. A distinction among the participants in the experimental group was seen as nearly half of them preferred to be offered both types of exam and the rest students desired no change. The students' scores displayed no significant change in relation to the test administration mode. In addition, it was found that both exam versions overlapped with respect to reliability and validity values, which showed that there was no obstacle to use computer assisted exams as an alternative to traditional way of assessment.

All in all, although it seems hard to make a radical shift from traditional testing to technology-based one, especially within the context of Turkey when the participants' attitude ratings are taken into consideration, students are capable enough in order to handle and use computer-assisted systems. Therefore, it is quite logical to utilize computers as testing tool alternatives and to make students acquainted with this smart system. Also, it is crucial to mention that computer-

assisted exams have no obstacle or drawbacks in term of offering reliable and valid tests, which are seen two backbones of tests, many advantages supplied by system to teachers, institutions and students and making students get equal and comparable scores to paper-based tests.

Keywords: computer-assisted assessment, web-based assessment, technology acceptance model, test-mode effect

Supervisor: Asst.Prof.Dr.Hüseyin ÖZ, Hacettepe University, Department of Foreign Language Education, Division of English Language Teaching



BİLGİSAYAR TEMELLİ ÖLÇME-DEĞERLENDİRMENİN İNGİLİZCE ÖĞRETMEN ADAYLARININ SINAV BAŞARISI VE TUTUMU ÜZERİNE ETKİSİ

Tuba ÖZTURAN

ÖZ

Ölçme ve değerlendirme eğitimin önemli bir parçasıdır, bu nedenle geçerli, güvenilir, kolay uygulanabilir testler hazırlamak, öğrencilerin kaygı düzeyini azaltacak, rahat sınav ortamları oluşturmak ve sınav kâğıtlarını tarafsız bir şekilde değerlendirmek oldukça önemlidir. Günümüzde bilgisayarların eğitim dünyasına kazandırılması ile ölçme ve değerlendirme alanında da yenilikler meydana gelmiştir. Örneğin, bilgisayar temelli ölçme-değerlendirme türleri birçok kurum ve öğretmen tarafından uygulanmaya başlanmıştır.

Bu çalışma, bilgisayar temelli ölçme-değerlendirmenin İngilizce öğretmen adaylarının başarı ve tutumu üzerindeki etkisini ortaya koymayı amaçlamaktadır. Bu kapsamda, Hacettepe Üniversitesi İngilizce Öğretmenliği bölümünde kayıtlı olan 97 öğretmen adayı bu çalışmaya katılmıştır. Öğretmen adayları deney ve kontrol grubu olarak ikiye ayrıldıktan sonra deney grubu vize sınavını bilgisayar ortamında kontrol grubu ise geleneksel yöntem ile almıştır. Ardından, katılımcıların tutumlarını öğrenebilmek için anket uygulanmıştır. Son olarak ise deney grubu arasından gönüllü olan kişiler ile röportaj yapılmıştır.

Sonuçlar deney grubunun bilgisayar temelli sınavlara karşı olumlu bakış sergilerken, kontrol grubunun bu tür sınavlara karşı ilgi göstermediğini ortaya koymuştur. Benzer bir şekilde, deney grubundaki öğrencilerin çoğu bilgisayar temelli sınavların, performanslarını artırdığını düşünmektedir. Ayrıca, deney grubu sınavları bilgisayar üzerinden almaları halinde, sistemi rahatlıkla kullanabileceklerini düşünmektedir.

Öğrencilerin bilgisayar temelli sınavlara karşı tutumları, cinsiyet, akademik başarı ortalaması, günlük bilgisayar/ internet kullanım süresi ve bilgisayar kullanım bilgisi gibi bazı demografik faktörlere göre değişiklik gösterebilir. Bu çalışmada, cinsiyetin herhangi bir bağlayıcı özelliği tespit edilmemiştir. Fakat bilgisayar kullanma açısından daha donanımlı olan öğrencilerin sistemi daha rahat kullandıkları ve

daha az kaygı duydukları görülmüştür. Ayrıca akademik başarı ortalaması daha yüksek olan öğrenciler sistemi daha kolay kullanmıştır.

Öğrencilerin sınav türü hakkındaki tercihleri ise şu şekildedir: Kontrol grubu geleneksel sınav yönteminin kullanılmasını tercih ederken, deney grubu bilgisayar temelli sınavların da geleneksel sınavlar ile birlikte sunulmasını tercih etmiştir. Katılımcıların sınav uygulama moduna göre puanları önemli bir değişiklik göstermemektedir. Ayrıca, her iki sınav türünde de geçerli ve güvenilir sınavların uygulanması mümkündür. Bu nedenle, bilgisayarların ölçme ve değerlendirme aracı olarak kullanılmasında herhangi bir engel yoktur.

Sonuç olarak, katılımcıların görüş ve önerileri dikkate alındığında, öğrencilerin bilgisayar temelli sınavları rahatlıkla kullanabilecek durumda olmalarına rağmen, ülkemizde sınav uygulama yöntemi açısından radikal bir değişiklik yapmak şuan için mümkün görülmemektedir. Bu nedenle, öğrencilerin bilgisayar temelli sınavlara alışabilmeleri için her iki sınav türü de öğrencilere sunulabilir.

Anahtar kelimeler: bilgisayar temelli ölçme-değerlendirme, internet temelli ölçme-değerlendirme, teknolojiyi kabullenme modeli, sınav uygulama yöntemi

Danışman: Yrd.Doç.Dr.Hüseyin ÖZ, Hacettepe Üniversitesi, Yabancı Diller Eğitimi ABD, İngilizce Öğretmenliği Bölümü

DECLARATION OF ETHICAL CONDUCT

I have prepared this thesis in accordance with the thesis writing rules and conventions of the Graduate School of Educational Sciences of Hacettepe University, and I hereby declare that:

- All the information and documents have been obtained on the basis of academic rules,
- All audio-visual and written information and results have been presented according to the rules of scientific standards,
- In case of using other works, related studies have been cited in accordance with the scientific standards,
- All cited studies have been fully referenced,
- I did not do any distortion in the data set,
- And any part of this thesis has not been presented as any other thesis study at this or any other university.

.....
Tuba ÖZTURAN

ACKNOWLEDGMENTS

It is very important to accomplish this thesis which has been mostly challenging for me and I would like to thank many people who support and guide me. Without them, it could have been impossible to end this work.

First and foremost, I am deeply grateful to my supervisor Asst. Prof. Dr. Hüseyin Öz for his endless guidance, support and patience during the writing of this thesis. I feel so lucky and privileged to be one of his students. He not only taught and guided me to gain precious things, but also motivated and encouraged me. Without him, I would not have achieved to write this thesis. I would like to thank him deeply from my heart for everything he did for me.

Also I would like to express my gratitude to the committee members Prof. Dr. Mehmet Demirezen, Prof. Dr. İsmail Hakkı Mirici, Assoc. Prof. Dr. İsmail Hakkı Erten and Assoc. Prof. Dr. Paşa Tefik Cephe.

I would like to express my sincere thanks and love to my friend Joanna Derman, who was always there to help me with the proof readings and to my colleagues/friends Ogün İter and Sevinç Akdoğan Öztürk for their guidance, support and suggestions. Besides, I am so thankful to Gülşah Uyar, my colleague, office mate and dear friend with her never-ending patience, help and motivation. I feel so lucky and happy to have them all in my life.

And special thanks to my parents, Ülkü and Mustafa Özturan, and to my brothers Ali and Kadir Özturan. With their endless patience, I could be able to end this thesis. My heartfelt gratitude goes to my best friends Güldeste Öztürk, Tuğba Dost and Sedef Örs who are like my sisters. They are always with me whenever I need them during this stressful and challenging journey.

Lastly, I would like to dedicate this thesis to the memory of my grandfather. I miss him so much.

TABLE OF CONTENTS

DECLARATION OF ETHICAL CONDUCT	viii
ACKNOWLEDGMENTS	ix
LIST OF TABLES	xiii
LIST OF ABBREVIATIONS	xv
1. INTRODUCTION	1
1.1. Introduction	1
1.2. Background of the Study	2
1.3. Statement of the Problem	5
1.4. Significance of the Study	6
1.5. Research Questions	7
1.6. Limitations of the Study	9
1.7. Assumptions of the Study	9
1.8. Operational Definitions of Terms	10
1.9. Conclusion	11
2. REVIEW OF LITERATURE	12
2.1. Introduction	12
2.2. Testing, Assessment and Teaching	12
2.3. Types of Tests	14
2.3.1. Knowledge Tests versus Performance Tests	15
2.3.2. Subjective Tests versus Objective Tests	16
2.3.3. Productive Tests versus Receptive Tests	16
2.3.4. Language Sub-skill Tests versus Communicative Tests	16
2.3.5. Criterion-referenced Tests versus Norm-referenced Tests	16
2.3.6. Discrete-point Tests versus Integrative Tests	17
2.3.7. Achievement Tests versus Proficiency Tests	17
2.3.8. Summative Tests versus Formative Tests	18
2.3.9. Diagnostic Tests versus Aptitude Tests	18
2.4. Principles of Assessment	18
2.4.1. Validity	19
2.4.2. Reliability	20
2.4.3. Practicality	22
2.4.4. Authenticity	22
2.4.5. Washback Effect	22
2.5. History of Computer Assisted Assessment	23
2.6. Computer Assisted Assessment Alternatives	25
2.6.1. Computer Based Testing	26
2.6.2. Web Based Testing	27
2.6.3. Computer Adaptive Testing	29
2.7. Benefits and Challenges of Computer Assisted Assessment	31
2.7.1. Benefits of Computer Assisted Assessment	31
2.7.2 Challenges of Computer Assisted Assessment	34
2.8. A Review of Related Studies Conducted in Turkey and Abroad	35
2.8.1. Abroad	35
2.8.2. Turkey	41
2.9. Conclusion	45

3. METHODOLOGY	46
3.1. Introduction	46
3.2. Research Design	46
3.2.1. Quantitative Research Design	47
3.2.2. Qualitative Research Design	48
3.3. Setting and Participants	48
3.4. Data Collection Instruments	52
3.4.1. Course Achievement Test	53
3.4.2. Questionnaire	56
3.4.3. Interview	60
3.5. Data Collection Procedure and Analysis	62
3.5.1. Data Collection Procedure	62
3.5.2. Data Analysis	65
3.6. Conclusion	69
4. RESULTS OF DATA ANALYSIS	71
4.1. Introduction	71
4.2. Quantitative Data Analysis	71
4.2.1. Results of Data Analysis for Research Question 1	71
4.2.2. Results of Data Analysis for Research Question 2	75
4.2.3. Results of Data Analysis for Research Question 3	77
4.2.4. Results of Data Analysis for Research Question 4	79
4.2.5. Results of Data Analysis for Research Question 5	79
4.3. Qualitative Data Analysis	80
4.3.1 Results of Data Analysis for Research Question 6	81
5. DISCUSSION OF THE FINDINGS	88
5.1. Introduction	88
5.2. Discussion of the Results	88
5.2.1. What are the prospective English teachers' attitudes towards computer-assisted assessment for their content knowledge course in general?	88
5.2.2. What is the relation between prospective English teachers' attitudes of computer assisted assessment and their demographic factors such as gender, GPA, daily use of computer and computer literacy?	90
5.2.3. Do the participants differ in their preferences of the exam mode?	90
5.2.4. Is there any statistically significant difference in mean scores between the students who took a computer assisted test (the experimental group) and those who took a paper based achievement test (the control group)?	92
5.2.5. Does the test administration mode influence the reliability and validity of the computer-assisted versus paper-based tests?	93
5.3. Conclusion	94
6. CONCLUSION	95
6.1. Introduction	95
6.2. Overview of the Study	95
6.3. Pedagogical Implications	98
6.4. Limitations of the Study	100

6.5. Suggestions for Further Research.....	100
6.6. Conclusion.....	101
REFERENCES.....	102
APPENDICES	111
APPENDIX 1: ETHICS COMMITTEE PERMISSION	112
APPENDIX 2: COMPUTER-ASSISTED ASSESSMENT QUESTIONNAIRE FOR THE EXPERIMENTAL GROUP	113
APPENDIX 3: COMPUTER-ASSISTED ASSESSMENT QUESTIONNAIRE FOR THE CONTROL GROUP	116
APPENDIX 4: INTERVIEW QUESTIONS	119
APPENDIX 5: COURSE ACHIEVEMENT TEST	120
APPENDIX 6: RELIABILITY COEFFICIENTS OF THE ITEMS IN COMPUTER-ASSISTED ASSESSMENT QUESTIONNAIRE FOR THE EXPERIMENTAL GROUP	129
APPENDIX 7: RELIABILITY COEFFICIENTS OF THE ITEMS IN COMPUTER-ASSISTED ASSESSMENT QUESTIONNAIRE FOR THE CONTROL GROUP	131
APPENDIX 8: ORIGINALITY REPORT.....	133
AUTHOR CV	134

LIST OF TABLES

Table 2.1. Contrasting Categories of ESL Tests	15
Table 3.1. Distribution of Participants by Gender	49
Table 3.2. Distribution of Participants by Exam-mode.....	49
Table 3.3 Distribution of Participants by GPA.....	50
Table 3.4. Distribution of Participants by Daily Use of Internet.....	50
Table 3.5. Distribution of Participants by Computer Literacy	51
Table 3.6 Independent Samples T-Test	52
Table 3.7 Distribution of Participants in Qualitative Study	52
Table 3.8. Split-half Coefficient of the Achievement Test in the Pilot Study	54
Table 3.9. Reliability Coefficients of Computer-assisted Assessment Questionnaire and Its Factors.....	59
Table 3.10 Reliability Coefficients of the Adapted Version of Computer-assisted Assessment Questionnaire and Its Factors	60
Table 3.11 Pearson Correlation Coefficients of the Questionnaires	60
Table 3.12 Normality Test of the Questionnaire	66
Table 3.13 Normality Test of the Achievement Test.....	67
Table 4.1. Independent Samples T-Test for Factor 1 – Intention	72
Table 4.2. Independent Samples T-Test for Factor 2 – Perceived Usefulness	72
Table 4.3. Independent Samples T-Test for Factor 3 – Perceived Ease of Use...	73
Table 4.4. Independent Samples T-Test for Factor 4 – Anxiety	74
Table 4.5. Independent Samples T-Test for Factor 5 – Computer Attitude	75
Table 4.6. Correlation Coefficients among Variables	77
Table 4.7. Chi-Square Test Results	78
Table 4.8. Crosstabulation of the Students' Exam Preferences	78
Table 4.9. Mann Whitney U Test Results	79
Table 4.10. Reliability Coefficients of Computer-assisted and Paper-based Tests	80
Table 4.11. Spearman's Rank-order Correlation Results	80

LIST OF FIGURES

Figure 2.1. The Relationship among Test, Measurement, Assessment, Teaching and Evaluation (Brown and Abeywickrama, 2010)	13
Figure 3.1. The Flowchart of the Research Model.....	47
Figure 3.2. A Sample Item on the System	55
Figure 3.3. A Sample Item on the System	55
Figure 3.4. Log-in Screen	56
Figure 3.5. Technology Acceptance Model	57
Figure 3.6. Histogram for Normality Test of the Questionnaire.....	67
Figure 3.7. Normal Q-Q Plot for Normality Test of the Questionnaire	67
Figure 3.8. Histogram for Normality Test of the Achievement Test	68
Figure 3.9. Normal Q-Q Plot for Normality Test of the Achievement Test	68
Figure 4.1. Comments of the Students.....	82

LIST OF ABBREVIATIONS

ANX: Anxiety

CA: Computer Attitude

CAT: Computer Adaptive Testing

CBT: Computer Based Testing

INT: Intention

PBT: Paper Based Testing

PEoU: Perceived Ease of Use

PU: Perceived Usefulness

TAM: Technology Acceptance Model

WBT: Web Based Testing

1. INTRODUCTION

1.1. Introduction

Assessment is an indispensable part of education because it not only gives feedback about the effectiveness of teaching and learning process but also serves as a tool for administrative, instructional and research purposes (Cohen, 2001). Given the importance of assessment in education, educators need to pay particular attention to some criteria like validity, reliability and the purpose for an effective assessment of teaching and learning outcomes. First and foremost, they need to define the purpose of assessment: *Do they assess students for “learning” or just use it for grading them? Should it be formative or summative assessment?* These are crucial points since there are different ways of assessment such as observation, peer-review, self-assessment, portfolio, tests and the like, and each of them may aid teachers for a different type of assessment. As Hughes (2003, pp. 33-34) emphasizes, an effective assessment includes both validity, which identifies whether the test actually measures what it is supposed to measure, and reliability, which refers to the consistency level of a test or a measurement instrument done at a different time and place. Certainly, it is equally important to prepare such measurement instruments, to provide a relaxed classroom atmosphere for test-takers and to grade them in an objective way. Nevertheless, assessment may be viewed as a challenging task and even as a burden sometimes for teachers and teacher educators.

With the introduction of computers into education, traditional assessment techniques are gradually being improved, or even replaced by techniques depending more on technology. In the past, technology was focused almost only in math and science classes, but today technology has also begun providing benefits to foreign/second language teaching and learning. At present, technology provides a number of opportunities in the assessment of learning and teaching outcomes in education. One recent popular computer-based technology that has emerged contributes a lot to assessment. As part of Information and Communication Technologies (ICT), computer based assessment methods have recently enhanced the assessment techniques commonly known as Computer-Assisted Assessment (CAA), Computer-Based Testing (CBT), Web-Based Testing (WBT)

and Computer-Adaptive Testing (CAT) (Bull and McKenna, 2004, p. 1). With its advantages of speed, flexibility and efficiency, this new mode of testing has led to innovative assessment practices as an alternative to traditional testing known as Paper-Based Testing (PBT) or Paper-and-Pencil Testing (PPT).

CAA now enables educators to improve their productivity and lead to innovation in their fields by standardizing test administration conditions. Whatever the size of the test population, CAA gives us the opportunity to set the same assessment conditions for the entire test takers (Bull and McKenna, 2004, pp. 8-9). Furthermore, it helps us overcome the challenges involved in test security since it is possible to store questions and answers in encrypted databases and to create randomized questions and answers from large question pools. Finally, CAA also has such great advantages like offering different test formats and the immediate presentation of different types of feedback, either to students or to test developers.

Given the importance of technology in foreign/second language education, the general aim of the present study was to investigate the impact of computer assisted assessment on the perceptions and course achievement performance of ELT teacher trainees at Hacettepe University. In the forthcoming sections, the background of the study, statement of the problem, significance of the study, research questions, assumptions and limitations of the study are presented to generate the overall structure of the present thesis. The last section of this chapter provides the operational definitions of the terms used in the study with a short conclusion paragraph.

1.2. Background of the Study

Technology is improving day by day and state-of-the-art technology has brought new trends for the realm of education, and computers have top-priority among the improvements because of the fact that they can address to learners who have different learning styles. Also, almost all learners have computer literacy to some degree, and they spend most of their time with an eye to computers. In this respect, some scholars assert that integrating technology and teaching/learning process trigger the success. For example, Sokolik (2001) and Salaberry (2001) support this assertion by pointing out that good teaching depends on both pedagogy and particular form of computer technology. In addition, Garcia and

Arias (2000) hold that computers are of importance to make learners gain an important skill: learning to learn. Lastly, Freeman and Anderson (2011, p. 199) highlight the importance of technology in language education by saying:

Technology is no longer simply contributing machinery or making authentic material or more resources available that teachers can use; it also provides learners with greater access to the target language. As a result, it has the potential to change where and when learning takes place.

With an overview of the history and place of computers in education, it should be stated that they have been used as a tool while teaching approximately for a long time, but especially from 1970s and onwards, so many researchers have turned their ways to determine the efficacy of computer-based instruction and computer-adaptive instruction (henceforth CAI) over students' learning (Ekane and Maiken, 1997; Sokolik, 2001; Warschauer and Healey, 1998). During this time, many improvements have occurred and brought about significant changes over teaching methods. Certainly all these improvements do not only include teaching; today learners can be assessed by computer technology as well. Specially designed programs help educators assess the learners' or test takers' knowledge on the subject. Named after as "computer-based testing", "computer-adaptive testing" or "web-based testing", the cutting-edge testing way offers many benefits for both educators and test-takers.

Computers have been utilized as a tool for testing since Frederick Lord launched his research in the 1970s when he created the theoretical structure of computer based testing (Cheng, 2009; Dunkel, 1999). The history of CAA in the field of language education goes back to the 1980s. Language Testing Research Colloquium published a proceeding in 1985 and gave the necessary information about CAA such as the construction of item pool, application of the test and the like. Since then, fruitful results have been attained; the universities across the USA, Britain and the Netherlands started to use CAA (Cheng, 2009; Roever, 2001). After the 1990s, research on computer-based language testing (CBLT) went up and many universities and institutions started to do research on the application and efficacy of CBLT. For instance, Education Testing Service started to implement TOEFL (Test of English as a Foreign Language) as computer-based in 1998, and it became a web-based exam in 2006 (Cheng, 2009). Another example done on the web is DIALANG. It is a project funded by the Council of Europe, and it provides assessment in fourteen European languages (Cheng,

2009). Initiated in the 1980s, CAA has obviously gained a new form, that is, Web-based Testing (WBT).

Assessment is an ongoing process; teachers can always assess students. On the other hand, tests are done at a specified time. In other words, it can be stated that tests are sub-components of assessment (Brown, 2004, p. 4). There are different types of assessment (summative, formative, proficiency, diagnostic and so on). Depending on the aim of assessment, the most suitable type of test designs (multiple choices, short answer, essay writing, gap filling and the like) can be implemented. By all means, CAA offers any type of these aforementioned test and/or assessment designs. Dunkel (1999) and Roever (2001) discuss this aspect of CAA by concluding that CAA supplies various forms of test items and that it is therefore possible to utilize many of them up to the aim of the testing. Multiple choice items, cloze tests, essays with large texts, discourse questions and the like constitute some examples of test-items. In addition, assessing listening skill becomes much easier via CAA.

As a specific type, CAA offers many reasons why it should be chosen for assessment (Brown, 2004). Of all, item response theory (IRT) or latent trait theory partakes a lot. According to IRT of assessment, tests can pinpoint the test-takers' level. Put it another way, the test-taker can see only one question at a time while being tested through computers, and the next question is defined according to the response of the test-taker to the previous one. In the case of right option, the next question becomes harder; the vice versa is valid, too. The initial question is in the middle-level, neither so hard nor so easy. In this regard, the test-takers cannot be bothered with easy questions or harder questions beyond their level (Brown, 2004, p. 14; Roever, 2001, p. 87). This is definitely an important characteristic of CAA that is not found in a traditional test delivery mode. Freeman and Anderson (2011, p. 200) specify the importance of CALL and CBT or CAT as follows:

Technology also allows teaching to be tailored to the individual to a greater extent than is normally possible. A few computer-assisted language learning (CALL) programs can even adapt to diverse learners by analyzing their input and providing customized feedback and remedial exercises suited to their proficiency. There are also programs that feature CAT so that students respond to test questions at an appropriate level.

Another vital feature in favor of CAA includes pedagogical implication. The test-takers can get immediate feedback and their scores upon completion the test. In

this sense, they can see their wrong and missing answers and see the correct one for each question, so they can be motivated for future studies (Brown, 2000, p. 160). In traditional exams, test-takers can take feedback from their teachers as well, but it might take time to learn the scores or see the incorrect or missing items with their correct answers. However, giving feedback right after any mission completed contributes a lot to washback effect or namely to pedagogical impact. In addition to the pedagogical impact on test-takers, this feature of CAA presents a lot of benefits for teachers as well because scoring all exam papers takes much of the time, and sometimes subjectivity and personal factors like fatigue and sleeplessness may violate the validity of the measurement instrument. CAA offers a list of students' grades just after expiring the allotted time for the exam. Lastly, computer assisted tests have eliminated the challenges like the limitation of time and space; students do not need to be in a specified place at a certain time.

On these grounds, the profits to be gained from technology for both teaching and assessing students cannot be denied. Every teacher should be given the opportunity to more or less utilize technology in and outside the class. This assertion should also hold true for the introduction and integration of assessment through technology into the curriculum.

1.3. Statement of the Problem

With the advent of computers and widespread use of the internet, the state of education has changed to a great extent over the past two decades. Now, it is compulsory to benefit from the technological facilities to make students motivated towards learning. As for foreign language teaching and learning, technology enhances authentic language for language learners, which is the backbone of Communicative Era. Furthermore, since the mid-2000s there has been an increasing use of technology in educational assessment worldwide. ICT facilities have started to enhance the assessment of the teaching and learning processes. CAA is a relatively recent test delivery method, so there is not a great deal of scholarly literature related to its role in language education.

Today, both teaching/learning practices and assessment procedures can be conducted by means of technological facilities. As stated before, CAA offers great advantages as it is a time-saving tool, gives immediate scoring and feedback,

prevents some probable problems of printing and assures objective scoring. In this sense, universities, schools and institutions such as Cambridge University, Stanford University, MIT and New Jersey Public Schools etc. started to make use of this genius tool for assessment.

The impact of CAA on teaching and learning is relatively a new field of investigation in foreign/second language education. There are various small-scale articles in the educational press and newspapers, but they mostly give only anecdotal evidence and advice. In other words, there are not many studies conducted to shed light upon the impact and comparability of CAA and PBT in foreign language education. Furthermore, gender differences, Grade Point Average (GPA), prior knowledge of computer literacy, daily usage of computers are among the challenging questions that need to be answered. Whether there is any difference related to students' grades between PBT and CAA or not is unanswered. In this regard, this study aimed to search the aforementioned issues within the knowledge base course of pre-service English language teachers.

1.4. Significance of the Study

This study aimed to examine the possible differences between CAA version and PBT version of the same exam (same questions) and the impact of exam mode on the test-takers. Accordingly, the significance of the study is given below.

No matter which test design is used, testing is so important for both teachers and students that it should be based on the principles of assessment like validity, reliability, authenticity, practicality, washback effect (or backwash effect) and item analysis. These related terms are defined in depth in the following chapter; however, it must be noted that they are compulsory for an effective testing. When comparing computer-based version and paper-based version of the same test, the former one can be labeled as "much more effective" rather than paper-based version. As assessment is a key to measure the outcomes of educational endeavors, so it is important to define whether computer-based tasks are equivalent to paper-based tasks and what factors influence the use of these two media (Noyes and Garland, 2008, p. 1352; Ployhart *et al.*, 2003, p. 733).

Assessment serves in different ways, especially in language education. Specifically, it is threefold: for administrative purposes, for instructional purposes

and for research purposes (Cohen, 2001). For instance, language tests such as TOEFL (Test of English as a Foreign Language), IELTS (International English Language Testing System) and Turkish national foreign language proficiency exam (YDS) address to instrumental purposes as the test-takers can use the results of these aforementioned tests so as to find jobs, to get promoted and so on. Today, large-scale exams like TOEFL and IETLS are done via CAA and YDS has started to be implemented via CAA recently. Hence, getting students to become familiar with CAA and/or WBTs with classroom applications will be fruitful.

CAA (WBT and/or CAT) is both time and place saving. Test-takers can take part in exam anywhere and anytime. When the allotted time is over, the system closes. Hence, there should not be any judgment. Also, when students take the CAA exam, they can get immediate feedback for each question or for overall exam upon completing the related set. This occasion can motivate students. CAA offers immediate scores for test-takers as well. On the other hand, when the exam is done on paper, students have to wait for teachers to learn their scores and grades. Certainly, it removes the burden over teachers as well. In addition, subjectivity over scoring cannot occur as machines score each item.

An analysis of the nature of CAA is needed for future studies despite all existent advantages. As stated before, this kind of research, done with ELT students in Turkey, is so rare. This study aimed to shed light upon this point lacking in the literature. After tested via a computer assisted assessment program for their midterm exam in *Approaches to ELT*, all the students in both groups defined their opinion by means of questionnaire and interview. In this way, advantages, disadvantages, probable shortcomings of CAA could be easily analyzed, so programs or applications for future tests can be organized in a more reliable and valid way.

Lastly, by referring to the benefits of CAA, teachers can be encouraged to use it for their courses as some of them may not be prone to use technology either teaching or testing (sometimes neither teaching nor testing).

1.5. Research Questions

On the basis of the evidence currently available, it seems fair to suggest that computer assisted assessment (CAA) has a lot of advantages for both students

and teachers. However, our review of literature indicates that there are not adequate studies conducted to evaluate the impact of computer assisted assessment on student's perceptions and exam success in their classes. After a brief review of literature, it was concluded that this kind of studies that focused on the investigation the role of computer assisted assessment on student's perceptions and exam success in language teacher education settings in Turkey was not so common. Before deciding on whether to invest in computerized assessment systems, educators and policy makers need to understand how much this technology may contribute to their particular educational context and need to become familiar with the opinions of the test takers. The present study will enable Turkish ELT teacher training institutions to make informed decisions about whether to invest in this field of assessment and to better understand what they must do if they decide to make this commitment. Therefore, this study was designed to examine in depth the impact of computer assisted assessment and paper-based assessment on the perceptions and course achievement performance of ELT teacher trainees at a large state university, along with the two core principles of assessment, i.e. validity and reliability. This study also explored some test taker characteristics such as computer anxiety, computer attitude, intention, perceived ease of use and perceived usefulness, as a result, the overall validity of computer-based tests.

On the basis of the aforementioned central points, the present study was designed to address the following research questions:

1. What are the prospective English teachers' attitudes towards computer-assisted assessment for their content knowledge course in terms of
 - a) intention,
 - b) perceived usefulness,
 - c) perceived ease of use,
 - d) anxiety, and
 - e) computer attitude?
2. What is the relation between prospective English teachers' attitudes of computer assisted assessment and their demographic factors such as

gender, grade point average, daily use of internet/ computer, computer literacy?

3. Do the participants differ in their preferences of the exam mode?
4. Is there any statistically significant difference in mean scores between the students who took a computer assisted test (experimental group) and those who took a paper-based achievement test (control group)?
5. Does the test administration mode influence the reliability and validity of computer assisted versus paper based tests?
6. What are the general attitudes and opinions of the participants toward the computer-assisted and paper-based tests?

1.6. Limitations of the Study

Likewise any research project undertaken, this study was not without some limitations. Firstly, the study group covered 97 pre-service English language teachers studying at a state university in Turkey. Of these participants, only half of them took the test on computer. More participants can be included in future studies; in this sense, the findings related to the test-takers tested on computer will be better to generalize to a larger population in the context of Turkish education system. Secondly, computerized adaptive testing is the paradigm of the 21st century. Unfortunately, as the researcher collected data from an achievement test within a content knowledge course for the study, CAT could not be the focal point. In addition, the student teachers' second language was English and the midterm exam of the content knowledge course was done in English. Although the participants were prospective English language teachers and they took computer courses in their first year at the university, any missing knowledge in terms of English and/or the content knowledge course might prevent them from showing their capacities enough in the exam. And lastly, some probable technical problems (like power-cut, hardware problems) might decrease the quality of the study.

1.7. Assumptions of the Study

The main assumptions of the study can be summarized under three components: first of all, the participants included all prospective English language teachers taking *Approaches to ELT*. In order to carry out the study in a desired way, the

participants were divided into two groups randomly. Hence, it was assumed that both groups included participants who had similar computer knowledge and success in their courses. Secondly, the participants took the exam either on paper or on computer as an achievement test (mid-term exam) of content-knowledge course. Clearly, the participants did not take the exam just for the study, but they took the mid-term exam, and the study was done in relation to it. Hence, it was assumed that participants would take the exam willingly and would not spoil the study's purposes. As the third and last component, it can be said that a questionnaire was implemented upon completion the test. The participants were assumed to take part in this questionnaire willingly and to give clear results.

1.8. Operational Definitions of Terms

Assessment is an ongoing process that is mainly based on the relation between curriculum and the outcomes of the learning-teaching process (Kress, 2009, p. 20), so it gives feedback not only about learners' performances but also teachers' contributions to learning process (Madsen, 1983, p. 5). Generally, it is constituted with different types like summative assessment, formative assessment, diagnostic assessment and the like.

Tests are the sub-components of assessment; in other words, tests are "just a form of assessment" (Hughes, 2003, p. 5), and they can be done in many designs like multiple choice, gap filling, essay writing etc. according to the aim of assessment. As tests are designed to evaluate one person's ability in a specific domain, the aim of assessment decides the shape of tests (Brown, 2003, p. 3).

Paper-and-pencil based testing is the traditional way of testing. With this type, questions are presented on a paper, and test-takers are required to answer the questions in relation to the item type. However, in this kind of traditional way a test has "fixed forms and un-weighted scoring" (Ward, 2002, p. 37).

Computer-assisted assessment is done by using information technology while testing. It is also known as e-assessment. Computer-assisted assessment has started to emerge as an alternative to paper-based tests since 1980s, and it has moved tests a step further because of the fact that "computerization results in an explosion in the range of possibilities that can be considered in designing, administering and scoring a test" (Ward, 2002, p. 37).

Web-Based Testing is a specific type of computer-assisted assessment and can be delivered via web browser as its name utters. It has brought a new trend for assessment: test-takers can take part in exam anywhere and anytime. In this respect, web-based tests have great advantages over computer-based tests (Roever, 2001, pp. 90-91).

Computer-Adaptive Testing is the paradigm of the 21st century. According to it, each student is tested in relation to his/her performance during tests. Computer-adaptive tests utilize large item pools as each test starts with a question neither too hard nor too easy and goes on with a question depending upon the answers of previous questions; it tailors subsequent questions with respect to test-taker's answers. This can be achieved with item response theory which is also known as latent trait theory (Chapelle and Douglas, 2006, pp. 7-8).

Technology Acceptance Model is proposed so as to shed light upon the factors that affect users' acceptance of technology or the probable reasons behind users' refusal to technology. The model has some sub-scales such as intention, anxiety, computer attitude, perceived ease of use and perceived usefulness, and each of them determines external factors that affect internal beliefs (Davies *et al.*, 1989).

1.9. Conclusion

This part was devoted to give a frame of the study that was based on the technological tools to be used for assessment practices within the higher education context. After giving a short summary of related studies, the problematic areas on that issue and the importance of this kind of research were mentioned. Following to it, the next part will give a detailed explanation about testing, its principals and types as well as the brief history of technological testing tools, their alternatives and pros and cons.

2. REVIEW OF LITERATURE

2.1. Introduction

As an implementation of appraising educational activities in general, assessment provides deep insights into the effectiveness of instructional endeavors in all subject areas at all levels of education. Hence, teachers have to comply with the principles of assessment and administer the appropriate tests; in this regard, they can reveal the strengths and/or weaknesses of the students and of the teaching procedure. Also, the required materials for helping students improve can be provided. In recent years, with the improvements in technology, computers have been benefitted as testing tools since they offer some advantages for teachers, students, testing centers and institutions.

This chapter presents background to assessment in education in general and language teaching in particular. A conceptual framework is provided by touching upon the definition and principles of assessment, the history and types of computer assisted assessment, some benefits and challenges of computer assisted assessments and the related studies conducted in this field.

2.2. Testing, Assessment and Teaching

As Fisher and Frey (2007, p. 1) point out, “checking for understanding permeates the teaching world”. Assessment is a must at all levels of education as it offers a lot for better outcomes in teaching process. Therefore, teachers, especially language teachers, are supposed to have the necessary knowledge of assessment, its principles and test types so as to pinpoint the level of students’ foreign language use.

In general, ordinary people use the terms testing and assessment interchangeably. In educational contexts, however, they are different from each other. After learners participate in regular instructional activities, they are administered tests at regular intervals as part of the related curriculum. In this process, learners are always aware of the fact that their answers are measured and evaluated by their teachers or school administrators.

As for assessment that includes a variety of areas, it is done continually. Clearly, learners are always assessed. Consequently, it can be deduced that testing is not

equivalent of assessment, but a part of it (Brown, 2004; Brown and Abeywickrama, 2010; Hughes, 2003). More specifically, there is a difference between testing and assessment: testing is a way of scoring the outcomes of learning process by checking what have been learned or taught or achieved while assessment is done so as to provide a number of benefits for teachers and learners; for instance, learners' success or failure and teaching process effectiveness are determined at the end of the assessment procedure. Whether the required aims are gained by learners or not is checked with assessment. In addition, assessment is not a fixed domain; rather it may show difference from time to time and focuses on probable individual differences of learners (Madsen, 1983; Brown and Abeywickrama, 2010).

Assessment, in other words, is an essential part of teaching process and has a strong relation with testing as assessment “takes stock of what is needed, what one has to work with and how well one has done” mostly by means of tests (Bachman and Palmer, 2000, p. 71). In other words, “both testing and teaching are so closely interrelated that it is virtually impossible to work in either field without being constantly concerned with the other” (Heaton, 1988, p. 5). Brown and Abeywickrama (2010, p. 6) displays the relation among testing, measurement, assessment, teaching and evaluation as in Figure 2.1.

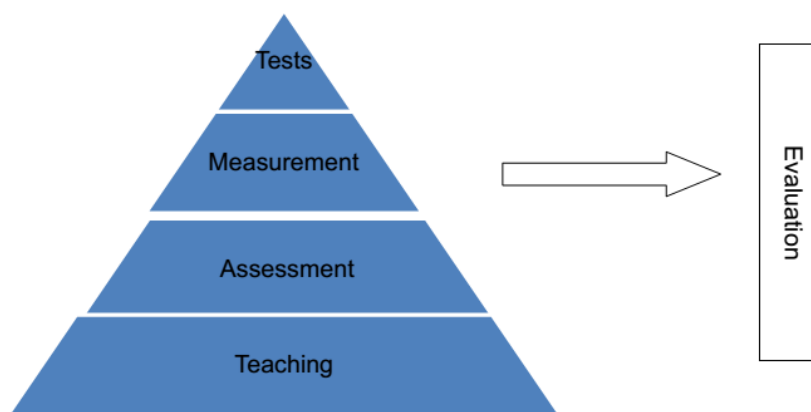


Figure 2.1. The Relationship among Test, Measurement, Assessment, Teaching and Evaluation (Brown and Abeywickrama, 2010)

As seen in the figure above, tests constitute the base for the teaching process. In line with the results of tests, a way of assessing students, teaching procedure can be carried out in a desired way or can be improved.

Assessment, which can be conducted in a variety of ways such as tests, self and peer assessment, portfolio, observation and so forth, is a vital phase in teaching/learning progress and in motivating students for further learning (Fisher and Frey, 2007; Gardner, 2012). These facts are ignored by virtually all the teachers. In fact, they should bear in mind the following points before preparing a test for an effective assessment procedure (Brown and Abeywickrama, 2010, pp. 52-53):

1. *What is the purpose of the test?* Why the test is generated constitutes the frame of the test's purpose. In this regard, the objectives of the test can be determined.
2. *What are the objectives of the test?* Test forms, functions and language abilities to be tested are defined in accordance with the test objectives.
3. *How will the test specifications reflect both the purpose and objectives?* While grading or generating a test, the test objectives should be ascertained by covering different skills.
4. *How will the test tasks be selected and the separate items arranged?* Practicality, content validity and reliable scoring should be enhanced.
5. *What kind of scoring, grading and/or feedback is expected?* Tests' purposes cause different test forms and the related feedback. Thus, how results are announced may vary in line with the purposes.

All these facts constitute the base for forming effective tests. Therefore, they should be taken into consideration all the time.

2.3. Types of Tests

According to Bachman and Palmer (2000, p. 8), "language tests can be a valuable tool for providing information that is relevant to several concerns in language teaching". Undoubtedly, there is not only a way of assessing students, but there are various kinds of tests in order to supplement an effective assessment. In this respect, teachers have to define the course objectives clearly, and then choose the most suitable way of testing as in general, different test models are needed for different sets of aims; that is to say, the aims may influence the shape of test

(Bachman and Palmer, 2000; Brown and Abeywickrama, 2010; Cohen, 2001; Hughes, 2003; McNamara, 2000).

Madsen (1983, p. 8) compiled and contrasted the test models specifically used in English language teaching as displayed in Table 2.1. Accordingly, each kind of test addresses to different teaching outcomes. For example, students can be assessed with formal tests at a specified time or they can be assessed during teaching hours informally. When speaking and writing skills are put on the front, productive tests are required. On the other hand, receptive skills, reading and listening, can be only assessed by means of receptive tests.

Table 2.1. Contrasting Categories of ESL Tests

Knowledge Tests	Performance (Skills) Tests
Subjective Tests	Objective Tests
Productive Tests	Receptive Tests
Language Sub-skill Tests	Communication Skills Tests
Norm-referenced Tests	Criterion-referenced Tests
Discrete-point Tests	Integrative Tests
Proficiency Tests	Achievement Tests
Summative Tests	Formative Tests
Diagnostic Tests	Aptitude Tests

2.3.1. Knowledge Tests versus Performance Tests

Learning outcomes have been commonly assessed with traditional tests including multiple-choice, gap-filling and/or true-false items for long years, which constitute the base for the knowledge-based tests. On the other hand, especially with the birth of communicative competence notion in the 1970s, a new aspect of assessment came to scene: performance-based assessment (Brown and Abeywickrama, 2010; Freeman and Anderson, 2011; Hymes, 1972). Accordingly, students' or language learners' productive and integrated skills should be of importance while they are assessed, which usually gives a way of authenticity in assessment. Performance-based assessment, contrary to knowledge-based tests, covers various types such as portfolios, journals, observations, self and peer assessment, conferences and interviews (Brown, 2004; Özmen and Balçıklanlı, 2012). When compared these two types of testing, it is fairly explicit that performance-based assessment, rather than knowledge-based one, can enhance direct testing and assure high content-validity of tests by utilizing authentic materials or tasks once students are assessed.

2.3.2. Subjective Tests versus Objective Tests

Not only item types but also the scoring of these items causes a distinction among tests. While teachers can evaluate some tests like multiple choice or true/false without any judgment, scorer or human effect may be in due for some tests like essay type or open-ended questions. Although objective tests overweigh in terms of evaluating students without fall, subjective tests require a natural production of language skills (Heaton, 1988; Hughes, 2003; Madsen, 1983; McNamara, 2000).

2.3.3. Productive Tests versus Receptive Tests

Language skills are divided into two categories as productive skills, including speaking and writing, and receptive skills, constituted with reading and listening. As each of these skills differs in their nature, different test models are needed when language learners are measured. Accordingly, students have to be active participants and produce a written or spoken act in productive tests contrary to receptive tests in which students simply read or listen to a text and answer the related questions (Heaton, 1988; Madsen, 1983).

2.3.4. Language Sub-skill Tests versus Communicative Tests

Apart from language skills consisting of receptive and productive skills, language areas like vocabulary, grammar and pronunciation sometimes need to be assessed isolated, and the most suitable way of doing so is generating language sub-skill tests (Madsen, 1983). As also indicated by Heaton (1988), the sub-skill tests show language learners' usage of target language that is related to the formal aspects of language. On the other hand, especially after the 1970s, the communicative approach played important role in foreign language teaching and testing. Accordingly, communicative tests display exactly how successful students are in using the target language for 'exchanging ideas and information' (Madsen, 1983, p. 9). In order to assess the language learners' communicative competence, real-world tasks and authenticity should be of important in designing communicative tests (McNamara, 2000, p. 17).

2.3.5. Criterion-referenced Tests versus Norm-referenced Tests

Tests vary in comparing test-takers' scores in accordance with certain criteria or average scores. In criterion-referenced tests, as its name suggests, there is a defined cut-off score in order to give feedback to test-takers, whether they pass or

not. Each test-taker is rated by taking into consideration the criterion and by ignoring how the other students achieve. On the other hand, norm-referenced tests range test-takers in accordance with mean, median and standard deviation of the exam. That is to say, each test-taker is compared with the other test-takers (Bachman and Palmer, 2000; Brown, 1996; McNamara, 2000). Between these two types of tests, criterion-referenced ones supply direct and meaningful results for test-takers since they display each student's language capability satisfactorily in contrary to norm-referenced tests (Hughes, 2003).

2.3.6. Discrete-point Tests versus Integrative Tests

Language tests can be administered by dividing language into its parts, especially grammar, vocabulary and pronunciation, which refers to discrete-point tests. In these tests, language is evaluated item by item at a specific time. As for integrative tests, language is seen as a whole unit and tested by combining different skills together and by creating a real life atmosphere in tests. Cloze tests or dictation are assumed as examples for integrative tests (Heaton, 1988; Madsen, 1983; McNamara, 2000).

2.3.7. Achievement Tests versus Proficiency Tests

Most of the time or in certain settings (like entering a university) students are required to take a language test so as to prove that they have sufficient knowledge of the target language. On that occasion, students take proficiency tests which assess the mastery of the target language and are not framed with any curriculum or syllabus (Brown and Abeywickrama, 2010; Davies, 1983; Hughes, 2003; Madsen, 1983). TOEFL (Test of English as a Foreign Language), administered by ETS (Educational Testing Service), is an example for proficiency tests (Brown, 2004, p. 45). Conversely, achievement tests are utilized with an aim to display how much each test-taker has achieved or learned in classroom. Namely, progress or learning outcomes are determined with achievement tests; therefore, these tests are mainly based on the units or syllabus, and they indicate the needs or deficient areas of the students. In this regard, future learning and teaching process can be planned in relation to the results of achievement tests (Davies, 1983; Madsen, 1983).

2.3.8. Summative Tests versus Formative Tests

Within a classroom setting, a language teacher may either assess the learners' knowledge so as to give feedback about their progress or measure their grasp of the related syllabus in order to decide whether they pass or fail. Under this circumstance, summative assessment and formative assessment are in discussion. Formative assessment, of which components are feed-up, feedback and feed-forward, is commonly used in language classrooms, especially for deciding future learning and teaching tasks and for detecting the ongoing improvement of the learners (Frey and Fisher, 2011, p. 4). Informal assessment such as observation and quizzes are some examples that can be done as formative assessment. Summative assessment, on the other hand, are intentionally done "to report what has been achieved rather than to help teaching and learning" (Harlen, 2012, p. 91). Therefore, summative assessment is administered generally when a course or unit is completed. Midterm exams, final exams, proficiency exams and periodic review tests can be given as examples for summative assessment (Brown and Abeywickrama, 2010, p. 7).

2.3.9. Diagnostic Tests versus Aptitude Tests

Diagnostic test is a part of formative assessment; hence, they are very similar in their nature. Students' strengths and/or weaknesses can be detected by utilizing diagnostic tests which commonly used for students who are about to fail. In addition, diagnostic tests are administered so as to test language skills in isolation (Hughes, 2003; Smith and Cumming, 2009). For instance, pronunciation tests may show which sounds are difficult for language learners, and they can be integrated into the curriculum (Brown and Abeywickrama, 2010). As for aptitude tests, they are used for measuring language learners' ability and capacity to learn that language (Davies, 1983). The Modern Language Aptitude Test (MLAT) and the Pimsleur Language Aptitude Battery, based on several English language tasks, are two standard examples for aptitude tests (Brown and Abeywickrama, 2010, p. 11).

2.4. Principles of Assessment

Defining the aims of tests and choosing the most suitable test type should be done before conducting a test. However, it is not enough in order to have an effective test. In this sense, educators have to first consider some specific principles.

Validity and reliability are foremost among these principles. In addition, a practical and an authentic assessment should be supplied. Lastly, assessment has impact on test-takers, which refers to washback effect. Now, detailed explanations of these terms, namely principles, are explored.

2.4.1. Validity

As the core of any assessment, validity refers to the relation between the aim and the form of the assessment. In other words, the tests measure what they are supposed to measure once the tests are valid (Dean, 1996; Madsen, 1983; Stobart, 2012). Gronlund (1998) defines validity as “the extent to which inferences made from assessment results are appropriate, meaningful and useful in terms of the purpose of the assessment” (p. 226).

There are different types of validity that have emerged in the field of educational studies. For example, *face validity*, the degree that a test includes what it purports to assess, can be yielded. In this context, a speaking exam done to measure test-takers’ pronunciation ability assures the concept of face validity; namely, it evaluates the skill being assessed (Hughes, 2003). When each test-taker is directly tested in relation to the act to be measured, then *content validity* is supplied. To achieve this sort of validity, if a test requires assessing the listening skill of a learner group, making them listen in the target language should be done (Brown and Abeywickrama, 2010; Dean, 1996; Hughes, 2003). Another type of validity is criterion-related; *criterion-related evidence* or *criterion-related validity* entails the comparison of two tests that have same criterion. Once the correlation coefficient of the tests determines the level, it shows that the tests overlap. As an example, in a classroom setting, a teacher may use of any standard test in order to assess grammatical knowledge of the students. Also, any commercially produced test can be benefitted in the same topic. If they both reveal the necessary degree of correlation coefficient, then criterion-related validity is supplied. *Concurrent validity* and *predictive validity* are similar, but they differ in their administration time. In the former one, the tests are done at the same time. In the latter one, tests are done at different times (Dean, 1996, p. 247-248). Generally, each theme in language teaching covers a different construct like “communicative competence”, “fluency”, “proficiency” and the like (Brown, 2004, p. 25). For instructors, this might seem hard to measure, and indeed, it is not

compulsory for classrooms. However, by taking into consideration whether a test really overlaps with its theoretical bases, then *construct validity* is achieved (Brown and Abeywickrama, 2010; Dean, 1996).

However, this following question still lingers: how can instructors create validity tests or increase the validity of tests? There are some tips available to them, documented in available academic literature: firstly, direct testing should be done whenever feasible, and explanations should be made clear. Secondly, scoring should be directly in relation to the targets of tests. Lastly, reliability has to be satisfied. Otherwise, validity cannot be assured (Hughes, 2003).

2.4.2. Reliability

Reliability is the degree to which a test measures a skill and/or knowledge consistently (Scheerens, Glas and Thomas, 2005, p. 93); therefore, the similar scores are commonly achieved on a reliable test once the same exam is administered on two different days. It is important to note that Brown and Abeywickrama (2010) and Hughes (2003) both emphasize that the interval between the administrations of two tests should be neither too long as students might learn new things nor too short as it might change students' ability to remember the exam questions. Once the test is reliable, the test-takers will get more or less the same score no matter when the test is administered, on a certain day or on coming days, and teachers have to prepare and administer reliable tests so as to obtain similar results from the same students, but at a different time (Hughes, 2003, p. 36; Madsen, 1983, p. 179).

There are three strategies to obtain test-reliability: the first appropriate strategy involves *test-retest reliability*. To achieve this type of reliability, the instructor must administer the same test twice to the same group of students. If similar results are obtained from both tests, the test is reliable. However, testing the same groups with the same questions for the second time is not desirable both for teachers and for students (Hughes, 2003, p. 39); hence, the second strategy to obtain reliability, *split-half reliability method*, is commonly used by teachers as it is quite easy to administer, and it is time-saving. The basic principal of this method is to split the test into two parts and to give scores for both halves. These two sets of scores are used to obtain test-reliability (Hughes 2003, p. 40). There is one more reliability

method, namely *parallel forms reliability*. As its name suggests, teachers striving for this type of reliability may choose to prepare two different forms of tests that are intended to be used interchangeably; hence, these kinds of tests have many points in common, and test takers are assumed to obtain similar results from these two parallel tests in case the test is reliable (Bachman and Palmer, 2000, p. 20).

Reliable tests give predictions about to what extent measurement-related factors may have impact upon test scores. These factors can be grouped in the following categories: test factors that refer to the clarity of instruction, items, paper-layout and the length of the test; situational factors that refer to the conditions of the room; and individual factors that cover the physical and psychological state of test-taker. Cohen (2001) states that all of these factors should be considered while interpreting the reliability of any test (p. 525).

Hughes (2003, pp. 44-50) provides instructors with some clues about how to create a reliable test:

1. Take enough samples of behavior
2. Exclude items which do not discriminate well between weaker and stronger students
3. Do not allow candidates too much freedom
4. Write unambiguous items
5. Provide clear and explicit instructions
6. Ensure that tests are well laid out and perfectly legible
7. Make candidates familiar with format and testing techniques
8. Provide uniform and non-distracting conditions of administration
9. Use items that permit scoring which is as objective as possible
10. Provide a detailed scoring key
11. Identify candidates by number not by name
12. Employ multiple, independent scoring

Teachers have to prepare reliable tests by taking account these clues, because once a test is not reliable, it is impossible to ensure the test's validity.

2.4.3. Practicality

Apart from core elements of a test, which are validity and reliability, a test has to be practical in order to be administered in an effective way. Hence, no test should be unnecessarily expensive, hard to be administered or time-consuming while being evaluated (Brown and Abeywickrama, 2010, p. 26).

2.4.4. Authenticity

Learning and teaching any foreign language cannot be considered separately from authenticity as it supplies real-world atmosphere for learners. In this way, target language can be taught by integrating its culture, and the same principle is valid for testing students in the target language. Not only teaching materials but also tests should be prepared in relation to authenticity. Accordingly, an authentic test's language needs to be natural, the test items should not be isolated and the topics should be meaningful and relevant to learners' interest (Brown and Abeywickrama, 2010, pp. 36-37).

2.4.5. Washback Effect

Undoubtedly assessment procedure is in line with the teaching and learning process; there may be a positive (beneficial) or negative (harmful) correlated relation between testing and teaching. Clearly, testing results certainly affect the pedagogical techniques in two ways, by supporting teaching process or by revealing some deficient sides of students, teachers and/or lessons, which is called as "washback effect" or "backwash effect" (Hughes, 2003, p. 1).

Teaching and testing are not separated from each other. It is quite explicit that in some situations, the teaching process is carried out fairly effectively. However, testing may be inadequate in some respects. On the other hand, sometimes teaching may not be appropriate, but testing is done in a well-planned way. Testing is not supposed to follow teaching, but the instructors can use testing to support teaching and to enhance its quality. For instance, students have a chance to benefit from a well-made test in any course, specifically English, by having motivation and enthusiasm to the course and by being aware of the points or language areas that need attention to master in it (Hughes, 2003; Madsen, 1983). However extensive and continuous testing may kill a student's love for learning. As an example, in China, there are two crucial high-stakes exams for students that

decide which upper-secondary schools and universities they will attend. In this respect, during many English classes, exams and exam practices are continually done, and unfortunately, this hinders students' desire to learn English within communicative frame, but they focus on exams (O'Hare, 2011, p. 132).

As washback effect has the capacity to improve the quality of education, assessment should be done in a well-planned way. Firstly, direct testing must be utilized in order to assess the language learners' skills in a desired way. For example, if a teacher aims to reveal the level of students writing, the assessment procedure must include writing tasks. Generally, language teachers may tend to assess grammar and/or vocabulary knowledge of students and this approach may obstruct the desired washback effect. Secondly, criterion-referenced assessment helps the students develop their knowledge without constantly concerning themselves over their peers' performance on tests. In this regard, a trustworthy picture of the teaching and learning process outcomes can be reached. Thirdly, when formal tests are administered, not only scores but also wrong items of each student should be shown with their correct answers so as to motivate them for further learning and to help them have a deep-insight over their success. In addition, positive reinforcement should be provided by praising students' correct answers (Brown and Abeywickrama, 2010; Hughes, 2003).

2.5. History of Computer Assisted Assessment

Integrating computer technology into the education dates back to the 1970s. Before that time, computers were so big that a single one could fill an entire room, so it was not convenient to incorporate them into academic settings – especially for those that were foreign language oriented- until later. Launched in the 60s at the University of Illinois, PLATO was a pioneer project that utilized computers for teaching and testing as well. In the 1970s, however, improvements in the 'miniaturization' of the computer components were done, and this enabled education to be integrated with technology (Sokolik, 2001, p. 477). Later, Frederick Lord created large-scale tests based on computers' capacities (Cheng, 2009, p. 1764). In the 1980s, computers became increasingly accepted by instructors and researchers as potential assessment tools. In other words, computer-assisted tests were utilized in certain fields like chemistry, mathematics, physics, engineering and so on though they were mainly administered in line with the

research aims (Carton *et al.*, 1991, p. 260). However, in the 1990s, advances in computer industry started to go up in education with the spread of World Wide Web (www), and then a variety of computer-based systems such as WebCT, Blackboard, Moodle, ATutor, Claroline, Desire2Learn and the like started to be used in education (Cisar *et al.*, 2010, p. 149). Namely, the rise of computer use in the teaching process and improvements in mental assessment theories led to precise, effective and individualized ways of assessment by making use of technology (Tung, 1986, p. 13). In this respect, national educational programs, especially in the USA, began to discover the effectiveness of computer assisted testing for students, teachers and institutions and to use computers as assessment tools as a part of their curricula. By the same token, Norway and Singapore launched to evaluate the pros and cons of the computer assisted assessment and to put forward the possible ways of utilizing computers in exams (Russell, Goldberg and O'Connor, 2003, p. 280).

Computer-based language testing (CBLT) studies occurred in the 1980s when two specific kinds of computer-assisted assessment, exam testing and class testing, were identified by Jones. Furthermore, Language Testing Research Colloquium compiled proceedings titled as "Technology and Language Testing" in 1985. These proceedings revealed that researches dealing with CBLT started to be done, and as a result of these studies, many institutions like universities and testing organizations benefitted from computers as assessment tools, because stakeholders in educational settings slowly gained the necessary knowledge about the construction of item-bank, the submission of tests and the application of adaptive tests. Starting in the 1990s, studies related to the computer-based language testing and submissions of this type of testing went up as exams could be done in an accurate and fast way (Akdemir and Oğuz, 2008). For instance, in 1998, computer-based TOEFL was started by the Educational Testing Service, and in 2006, the web-based version of the same exam was started to be administered, which constituted the base of the transition from paper-based tests to computerized tests and then to web-based tests around the world. The University of Cambridge Local Examinations Syndicate is another example for the institutions administering computer assisted tests (Cheng, 2009, pp. 1764-1765). In addition, the DIALANG project, funded by the Council of Europe, implements

computer-based language tests in fourteen European languages. It provides immediate feedback on scores and explanations and suggestions about the items in the test, which is a way of fostering the self-assessment levels of the test-takers. Also, the DIALANG is not limited with the types of items such as multiple-choice, gap-filling and the like, but offers at least eighteen different forms of questions (Alderson, 2000, p. 600).

In summary, computer assisted assessment (CAA) has become so widespread that almost all universities and schools abroad have made CAA a part of their curriculum. For example, the Massachusetts Institute of Technology, Cambridge University, Stanford University and Princeton University are the pioneer educational settings to implement computer assisted tests within their curriculum. Not only these institutions, but also many secondary schools, mainly in USA, have started to use CAA as each student can easily be pinpointed. Furthermore, CAA is now a practical concern among thousands of people all around the world. For example, high-stakes exams like TOEFL and IELTS have started to be done on computers, and in this way, authentic assessment (especially for listening skills) can be supplied. In addition, hundreds of thousand people around the globe can take the test without any time and place restriction.

2.6. Computer Assisted Assessment Alternatives

Computers are no longer passively involved devices in our lives; on the contrary, they take part in many different walks of life actively. For example, call centers, banks, companies and schools make use of computers as smart helpers. This role change in computer applications goes back to the late 1970s. Since that time, computers have had a vital place in the world, especially for educational purposes. In addition to the widespread use of web and computers as teaching sources both inside and outside the class (especially for distance education), computers have launched to offer testing alternatives for teachers as well. Today, it is estimated that nearly 1000 computer-assisted assessments are done each day in the UK (Lilley, Barker and Britton, 2004, p. 109). These assessment models do not only refer to the traditional tests that are administered on computers in class under the supervision of proctors. It has different sorts of alternatives which are named as computer-based testing (CBT), web-based testing (WBT) and computer-adaptive

testing (CAT). Now, the paper turns the way so as to explain each of them respectively.

2.6.1. Computer Based Testing

Computer-based testing refers to make use of computers while preparing questions, administering exams and scoring them (Chapelle, 2001), and with the advent of using computers as testing devices since the 1980s, a different point of view has been gained so that more authentic, cost-saving, scrutinized and controlled testing environment can be achieved, comparing to traditional paper-and-pencil based one (Linden, 2002; Parshall *et al.*, 2002; Wang, 2010; Ward, 2002). Although computer-based testing literally started in the late 1970s, or namely in the beginning of 1980s, it had been always thought as an alternative to paper-based testing before (Folk and Smith, 2002), because “one size fits all” solution across testing programs was not desired at all (Ward, 2002, p. 37).

Computers have brought many advantages. First of all, they have the potential to offer realistic test items like media, graphics, pictures, video and sound (Chapelle and Douglas, 2006, p. 9; Linden, 2002, p. 9). Therefore, students can be involved in a real-life testing environment where there are many integrated activities. In other words, students can respond to computers orally, draw on the screen while answering the question, see and interpret graphics or tables for an open-ended question and so on, and handicapped test-takers can take the exams on computer with great ease. CBT also supplies immediate feedback and scoring (Chapelle and Douglas, 2006; Parshall *et al.*, 2002), which has significant impact over pedagogy (test-takers can grasp their mistakes when immediate feedback is offered upon the completion of the test) and eases teachers' workload of scoring all papers – teachers may spend much time on scoring exam papers, and also, generally they cannot give enough feedback about each student's mistakes, or even if they provide feedback, it may be so late that students do not remember the questions or their answers. Another issue that should be mentioned here is that especially for open-ended questions, subjective-scoring may be in due. However, thanks to computer technology, objective scoring can be achieved, and handwritten problems are disappeared, too. And the last important feature of CBT or CAA is that the examiners can collect data about the exam such as how many questions have been answered correctly, how many of them have been omitted and how

many minutes have been spent for each question, which is called as *response latency* (Parshall *et al.*, 2002, p. 2).

Since the beginning of using computers as testing tools, many different computer based test delivery modes have come to scene by referring to different types of needs or to different approaches: computer-adaptive testing (CAT), linear-on-the-fly testing (LOFT) or computerized fixed tests (CFT), computerized mastery testing (CMT) (Ward, 2002, p. 38) and automated test assembly (ATA) (Parshall *et al.*, 2002, p. 11). CAT is totally performance or individual based testing. The more a candidate answers questions correctly, the more challenging questions appear on the screen, and vice versa. On the contrary, LOFT or CFT has fixed time and test-length for all test-takers. Exam security is the main goal in LOFT, rather than having psychometric values as in CAT (Parshall *et al.*, 2002). As for CMT, it aims to divide test-takers into mastery and non-mastery groups (Folk and Smith, 2002, pp. 49-50). Lastly, ATA chooses items from an item pool in regard to the test plan and makes use of content and statistical knowledge. This kind of test has fixed time and is not in adaptive mode (Parshall *et al.*, 2002, p. 10).

2.6.2. Web Based Testing

Greg Kearsley (1996) emphasized the importance of web and its future potential as an educational tool many years ago; and we have been experiencing it recently. Not only is Web a means of delivering information, material, news and so on from one part of the world to the whole, but also it is the most commonly used and significant benefit of teachers for a variety of things like searching different types of materials, teaching for distance education, presenting, preparing tests and delivering them (Khan, 1997).

For many years, visionaries and futurists have been telling us that one day we would have quick and easy online access to all of the world's information. Well, the future has arrived; it is called World Wide Web (WWW) and its growth in the past few years has been phenomenal. Already the web is showing us how global networks will transform education.

The reason lying behind this change is that since 1990s, international connectivity has not been limited only to teaching staff at universities and to their use of network in computer labs, and without any doubt, it has brought many differences. As for the testing applications, universal access to computer-assisted assessment has been introduced, and a bulk of opportunities for autonomous learning and self-

assessment has spread all around the world, and so have computer-based applications. Today, thanks to web-based applications, students and teachers can be universal and universally in touch (Chapelle, 2001, p. 23; Hudspeth, 1997, p. 353).

As an alternative of CAA, web-based testing is specifically driven and delivered by means of web, and it means that the tests can be taken anywhere and anytime, which constitutes the great advantage over traditional paper-based and computer-based tests (Nichols, 1997; Roever, 2001). Hudspeth (1997) states many reasons to conduct tests on the web: *“the highly interactive nature of the medium, the volatility and currency of information, remote access of instruction and instructional resources”* (p. 356). The web system, also, makes it possible to create unique exams, and it is based on an important mathematical content (McGough et al., 2001).

In three phases (preparation, delivery and assessment), a question can be created on the web. Accordingly, an item is on the threshold of being created at *authoring time*. Teachers can prepare questions and store them in an item bank by using web tools. Then, questions or items are selected in order to conduct the test. The selection of the items is done either statistically by teachers themselves or dynamically by the system at run time (Brusikolovsky and Miller; 1999, p. 2). After delivering the items and conducting the exam, examinees' answers are assessed as correct, incorrect or partially correct. On the web technology, preparing, delivering and assessing questions are based on HTML code (Brusikolovsky and Miller, 1999, pp. 2-3).

Roever (2001, pp. 90-91) holds that WBT is threefold as low-stakes assessment, medium-stakes assessment and high-stakes assessment. Low-stakes tests are used to give feedback about examinees' performances over a certain subject or skill. The examinees can take these tests wherever they want. On the other hand, medium-stakes assessment covers midterm and final exams done in classes, placement tests or any tests that have impact on the examinees' lives. These kinds of tests are carried out by proctors in a lab. And lastly, high-stakes assessment is the one of which results may affect greatly the examinee's life like being accepted to a university or certification programs or citizenship tests and so

on. Among these three types, WBT is much more useful when it is done for low-stakes assessment.

Many schools and universities have started to benefit from web technology while administering exams. One of them is Iowa State University that has created the WebCT. This smart system does not require any technical information so as to use it, and teachers can easily create and publish online courses and exams (Chapelle and Douglas, 2006, p. 63). Apart from the WebCT, Hot Potatoes, Discovery School Quiz Center, Blackboard and Questionmark are free accessible online tools to be utilized (Chapelle and Douglas, 2006, pp. 72-73).

2.6.3. Computer Adaptive Testing

The cutting-edge technology in computer applications in education has given way to assess test-takers' knowledge and ability by computers and an assessment theory which is known as Item Response Theory (henceforth IRT) (Stevenson and Gross, 1991, p. 224). This new assessment model is computer-adaptive testing that is based on each student's performance during the exam, and definitely, this new individualized exam model (Wainer, 2000, p. 1) offers more confidential testing atmosphere for both teachers and students (Guzman and Conejo, 2005; Linden and Glas, 2002).

The last mode of CAA, computer-adaptive testing (CAT), has been utilized for many years. The cycle of CAT begins with a question that is neither so easy nor so difficult. According to the answer of each test-taker to the item, which question to be asked from the item pool is decided. More clearly, if a test-taker answers a question correctly, the next one will be harder or on equal difficulty. On the contrary, if a test-taker answers a question incorrectly, the next one will be easier. Hence, CAT is said to be based on performance (Chapelle, 2001; Flaugher, 2000; Guzman and Conejo, 2005; Lilley, Barker and Britton 2004). Students can see each item on screen at a time, and they cannot skip the questions. While the test-takers are busy with each question, the system calculates the scores and decides which question will be next in relation to the previous answers given by the test-takers (Brown, 2004; Henning, 1991; Hughes, 2003). This measurement model in CATs is known as Item Response Theory or Latent Trait Theory of which

mathematical bases were outlined by Lord and Novick around 1970s (Tung, 1986, pp. 4-5).

The idea lying behind IRT goes back to the psychological measurement model, put forward by Alfred Binet and today known as the Stanford-Binet IQ test (Linden and Glas, 2002). Binet's idea of measuring each test-taker separately and according to their performance while they are taking the test has been accepted as the only adaptive testing approach for more than fifty years (Cisar *et al.* 2010; Reckase, 1989), but there was one drawback stated about this smart system: despite its truly adaptive side, experienced and skilled teachers (examiners) might be needed in order to administer large-scale tests. Therefore, it was practical only for small-scale tests (Madsen, 1991). However, the US military (the US Department of Defense with its Armed Services Vocational Aptitude Battery - ASVAB) (Cisar *et al.*, 2010; Reckase, 1989; Wainer, 2000) started to carry out research on the effectiveness of CAT and IRT principles after the launch of using powerful computers in the late of 1970s, and the first CAT was started to be implemented in the mid of 1980s. Then, it spread to be used for other departments and in other institutions like the National Council of State Boards of Nursing and the Graduate Record Examination (Linden and Glas, 2002, p. 8). Today, CAT is used not only for small-scale exams but also for large-scale high-stakes exams as well. For example, Graduate Management Admission Test, Microsoft Certified Professional and Test of English as a Foreign Language have been shifted from CAA to CAT (Lilley, Barker and Britton, 2004, p. 110), and SIETTE is a web-based computer adaptive testing system used in Spain (Guzman and Conejo, 2005, p. 688). In addition, IRT has many different models in it like "the one-parameter, two-parameter and three-parameter logistics models" that are mostly known and used models today (Henning, 1991, p. 212; Wainer and Mislevy, 2000). Each of them measures different amount of knowledge and different number of people at a time. For instance, as the name suggests, one-parameter model calculates test-takers' performance on a single latent ability and difficulty continuum. On the other hand, two-parameter model also estimates discrimination continuum. As for three-parameter model, it integrates ability-difficulty and discrimination continua. Moreover, the number of people attending tests differs in each parameter. One-parameter model calculates people between 100 and 200 while two-parameter

model used 400 or more people for calculation. Three-parameter model needs 2000 or more people so as to make the required calculation (Henning, 1991, p. 212; Wainer and Mislevy, 2000).

2.7. Benefits and Challenges of Computer Assisted Assessment

Computers have been commonly used as assessment tools in recent years; it has become a new trend into educational world. Undoubtedly, this is because it has brought many advantages and improvements for both educators and test-takers. However, no technological invention or improvement has benefits only. This part is devoted to explaining pros and cons of using computers as assessment tools.

2.7.1. Benefits of Computer Assisted Assessment

Computer-assisted testing or computer-adaptive testing offers a number of benefits for assessment. First and foremost, the efficacy of testing has gone up thanks to computers. It is an undeniable fact that the virtual learners in the 21st century are a lot keen on technological tools. In this regard, teaching methods have started to be integrated with computers and mobile phones. By the same token, learning on the move has become the aim of many teachers at all levels of education. This is not only valid for teaching but also covers testing issues, namely, “digital natives” can present better results if they are tested with computers (Sokolik, 2001).

Most probably one of the most vital advantages of CAA, or WBT, is the flexibility in terms of time and place, especially for large-scale tests. Internet lets examinees take the exam wherever they are and whenever they want (Alderson, 2000; Noyes and Garland, 2008; Roever, 2001; Vos and Hacquebord, 1998). Therefore, any probable problems with laboratory setting can be removed (Noyes and Garland, 2008, pp. 1362-1369). Test designers, also, can use the advantage of place flexibility by sharing tests with colleagues around the globe, and so, they can get feedback from them if necessary (Roever, 2001, p. 88).

Additionally, computers are the blessing benefit so as to offer very rich test content, and undoubtedly, this is the basis of communicative language testing (Brown and Abeywickrama, 2010; Choi, Kim and Boo, 2003; Noyes and Garland, 2008). In other words, visual cues like graphics, pictures, charts, and so on that are presented during listening tests (Laborda, 2010) or reading tests make the

exam atmosphere much more authentic, which is the key component of communicative language testing. However, Laborda *et al.* (2014) put forward the idea of using semi-directed communication tasks for speaking exams, because speaking exams that are totally based on computers do not look like so much communicative. Also, thanks to computers, exams are done in a standardized way, which means that each and every-test taker can have the advantage of participating in exam in the same way.

Another advantage of using computers as testing devices is related to Krashen's I+ 1 theory (Krashen, 1982) and Vygotsky's Zone of Proximal Development (Vygotsky, 1978). Accordingly, once a learner is taught beyond his/her current level, teaching is done more effectively, and without any doubt, teaching and assessment procedures cannot be thought separately.

The latest improvements in computer assisted assessment have given way to measuring each examinee's level separately: computer adaptive testing, which is driven by Item Response Theory (Drasgow, 2002, p. 4), starts the exam with a question neither so easy nor so difficult, and then, chooses next items depending on examinee's success or failure for each item (Alderson, 2000; Cohen, 2001, p. 521). Thus, computer adaptive testing supports Krashen's theory, which gives way to assess each learner's knowledge not with same questions for all, but with different questions depending on large amount of information (Sokolik, 2001).

The administrative aspect of CAA reveals another advantage: immediate test scores and feedback (Alderson, 2000; Cohen, 2001; Noijons, 2013; Yunxiang *et al.*, 2010). Contrary to paper-and-pencil based testing, which requires a long time from teachers to announce results, and most of the time, it is impossible for teachers to give enough feedback for each learner about their mistakes on exam items, computer-based testing makes delivering test scores and giving feedback just upon completing the test possible (Alderson, 2000, p. 595). Before administering any exam, a teacher can design the system about giving feedback in two different ways: feedback is given either after each item is answered or after the whole test is done. Both ways work a lot for learners as it is commonly stated that giving feedback right after any mission done has a crucial and meaningful impact on learning, which is useful for pedagogical purposes (Roever, 2001, p. 85), rather than postponing feedback as test-takers may forget exam items and

their responses because of the long time interval on that occasion (Alderson, 2000, p. 597). Furthermore, CAA allows test-takers to have a second chance for responding to a question after giving feedback. Clearly, once a question is not answered correctly, the system provides feedback and gives a second chance to the test-takers. Here may arise a question: which response given, before or after the feedback, should be taken into consideration for scoring? According to Callı *et al.* (2003), the second response should be scored, because learning has occurred after getting feedback. Still, this point needs to be shed light on through further studies. Davies (1983, p. 41), on the other hand, points to another advantage of giving feedback after exam: “students can see suggested course activities as a result of exam evaluation”, and without question, this leads to pedagogical improvements, and Robinson (1991, p. 159) supports this assertion by emphasizing the fact that CAA encourages learners to find their own learning strategies and to improve themselves at their own paces. Lastly, Choi, Kim and Boo (2003, p. 296) state that using computers as testing devices and giving feedback on the systems contribute a lot to washback effect of testing issues and improve construct validity.

Not only does CAA make test administration procedure easier, but also it increases the accuracy of scoring. By defining possible answers to the system, teachers erase the human error factor on scoring, and all examinees’ answers are measured objectively. Handwritten problem, on the other hand, is taken away with CAA (Madsen, 1991, pp. 240-241; Noyes and Garland, 2008, p. 1369). Moreover, computer technology decreases cost issues related to printing and delivering exams, especially for large-scale exams (Russell, Goldberg and O’Connor, 2003, p. 279; Vos and Hacquebord, 1998, p. 109).

Computer technology helps language teachers prepare tests within “an integrated and construct based approach, which is Portal” (Cheng, 2009, p. 1765). Portal uses computers in order to supply salient information about examinees’ performances on language by linking test designs and their skills. The reason lying behind this is the developments in computer based language testing. In this way, more arduous tasks can be given on exams. Cheng (2009, p. 1766) points out that “simulations with identified characteristics allow relevant features to be

manipulated in order to assess language learners' ability levels that are aimed to test."

Today special software programs such as Questionmark and Authorware are available on the market so as to aid test design, item editing, piloting and having an item pool. Moreover, these systems are greatly secured, because randomly assigned items constitute different bulks of tests (Alderson, 2000, pp. 595-596).

Despite all the advantages, there are some drawbacks of using computers while testing. Now the paper turns the way to explain possible challenges in depth.

2.7.2 Challenges of Computer Assisted Assessment

Computer assisted testing shows some drawbacks although it has so many benefits, and the most important drawback is about scoring. Test-designers can upload possible answers to the system, and the testing software program does not accept any different answers rather than the ones that have been defined in advance. Hence, test designers have to pay full attention to possible answers while defining them to the testing software before exam. Walker and White (2013, p. 135) give an example on that issue in their book:

"...For example, suppose a test includes a short-answer question where the correct answer is 'seagull'. The test could allow for 'gull' as an alternative answer... However, the student might give an answer such as 'hearing gull', 'black-headed gull', or even 'larus'. These answers might be correct, but if designer had not predicted them, the testing software would not accept them."

Another point related to it is that machine-based scoring limits the use of different types of questions, especially those of essays (Alderson, 2000; Luecht, 2001; Walker and White, 2013). Multiple choice, gap-filling, short-answer questions or cloze tests are easily used and scored while scoring essays is difficult since cohesion, grammar and vocabulary use, relevance and test structure are only taken into consideration by humans, not by computers (Cambridge ESOL, 2008).

It is quite explicit that a minimum degree of computer literacy is needed in order to take computer-assisted tests. In addition, reading questions or texts on screen is not same with reading on paper. However, Educational Testing Service made a noteworthy research about the effect of computer literacy over TOEFL success, and the results revealed that there was no difference between examinees who had computer literacy and those who did not have (Kirsch *et al.*, 1998; Roever, 2001; Taylor *et al.*, 1998).

Flexibility in terms of place and time contributes to the advantages of CAA while it may show security problems as well, because examinees are supervised by no proctors during the exam. More openly, whether or not an examinee uses any illegal means in order to answer questions cannot be specified whole (Noyes and Garland, 2008; Wainer and Eignor, 2000). Neil White and his team at Southampton University, however, used fingerprint identity so as to prevent any security problems. In this regard, each and every examinee can be identified with their use of keyboard and mouse (Semerci and Bektaş, 2005). Digital tools are easy to use for recording speech, and they are even easy to be deleted or changed or edited, too. This creates pitfalls for exam security (Roever, 2001; Walker and White, 2013). In addition, testing software systems are not capable enough of assessing the speaking part (Walker and White, 2013, p. 135).

Bridgeman (2009) and Noyes and Garland (2008) define computer hardware and software problems as breakdown, back-up and so on. In this case, exam-time can be wasted if computers are needed to be restarted, or sometimes, all responses given by examinees can be lost.

2.8. A Review of Related Studies Conducted in Turkey and Abroad

Educational institutions' decision to incorporate improved technologies into their classrooms has given teachers the opportunity to enhance their pedagogies and better measure their students' progress, and CAA applications, an umbrella term for CAT and WBT, are the latest ones that have gained popularity amongst many researchers and educators who are now looking to use CAA in a variety of contexts. The following analysis aims to introduce a series of related education technology studies that have been conducted abroad and in Turkey.

2.8.1. Abroad

Since the beginning of CAA's application, many researchers have started to search the topic by asking different questions. Now the paper will touch upon these studies by stating their aims and findings.

Whether an examinee shows similar scores in both exam versions or not has been the concern of many studies till today. Firstly, the American Psychological Association (APA, 1986) demonstrated that the participants, who had taken two different modes of the Graduate Record Exam, one being computer-assisted

exam, and the other one paper-based, had greatly equal scores. Same results were detected in the study by Parshall and Kromrey as well (1993). In a similar study, Bugbee (1996) compared these two different exam modes. Russel, Golberg and O'Connor (2003) emphasized the equivalence of CAA and PBT in their research. Higgings, Russell, and Hoffmann (2005) conducted a survey with 219 4th grade students in an attempt to define any probable score differences in reading comprehension between groups, resulting from the test-mode effect; their research revealed no statistically significant differences. Wang *et al.* (2008) investigated the comparability of test modes; thus, the related literature specifically done on this topic was analyzed by the authors. Finally, they stated that "administration mode had no statistically significant effect on K-12 student reading achievement scores" (p. 5).

Another survey, carried out by Stevenson and Gross (1991) in order to reveal the equivalence of test modes, revealed similar results about the CAT and the PBT. Their study was done at the Montgomery Country Public Schools and made use of a paper-based MEC (Minimum English Competency) exam, which was prepared and validated so as to prove English proficiency at those schools. It had a range of different series based on listening, speaking, reading and writing skills. In order to address to this study's aims, a CAT version of the MEC exam was developed by using 170 exam items in the item pool (100 items for reading part and 70 items for writing part). In total, 125 students took the paper-based MEC exam in December, and then, the same students took the same test, but on computer in March. The correlation between reading scores was 0.61 and correlation between writing scores was 0.66, which revealed the tight relation. Moreover, the authors mentioned that students enjoyed doing the computer-assisted MEC exam.

Choi, Kim and Boo (2003) carried out a survey with students from five different universities. Participants were randomly assigned to two different groups, and Group A firstly took PBT and then CAA while Group B took CAA firstly and then PBT. The exam items were for measuring listening and reading skills. At the end of the study, it was found that listening scores were higher on CAA; on the other hand, students got higher reading scores on PBT. Similarly, in some related literature, it was recommended that reading on screen was slower than on paper,

and the accuracy of reading went up, when students took their tests on paper (Creed *et al.*, 1987; Dillon, 1994; Muter *et al.*, 1982).

Chua (2012a, 2012b) made a survey in order to check the computers as testing tools, comparing to paper-based tests, too. In total, 140 student teachers participated in the study, and their test performances (time and scores), their motivation towards taking tests on computers and the validity and reliability of both test modes were analyzed. At the end, it was concluded that CAA revealed more reliable and efficient testing environment for prospective teachers. Retnawati (2015) compared the scores of the participants who took paper-based Test of English Proficiency with the ones who took computer-based version of the test as well, and the results revealed that scores in both exam modes were quite similar. Lastly, Logan (2015) aimed to search the students' performance differences up to exam administration mode within the frame of mathematics course. In total, 807 6th grade Singaporean students took the mathematics test with 24-item and the paper folding test either on computer or on paper. The results displayed that there was no significant difference; clearly, though the significant values were so close to each other, the scores gotten from paper-based tests were higher compared to the scores taken from computer-based tests.

Madsen (1991) touched upon a different perspective in the study; that's to say, he made a survey so as to evaluate the efficiency of CAT. In Madsen's study, 17 male and 25 female students (42 students in total), coming from a wide range of countries, including Thailand, Spain, Japan, Indonesia, Portugal, China and Vietnam, took a computer-adaptive reading test. Some of those students had little computer literacy, and none of them took computer-assisted tasks before. Upon the administration of CAT, students were asked their easiness with reading on computer screen. Apart from half of the Spanish students, almost all the students were happy with reading on screen. In addition, participants' feelings about being tested with computers were asked, and it was seen that prior experience with computers had no effect over test anxiety. In short, many students expressed their positive feelings about CAT.

Similarly, in Spain, Conejo *et al.* (2004) developed a system, called SIETTE (p. 1) and based on CAT and Item Response Theory theoretical background. With this smart system, teachers all around the world can control test-designs and exam,

and students can take the exams without any time and place restriction as well. In a similar point of view, Jo Ann Rowe (2007), a nurse educator in Toronto, Canada, compiled the related literature over that topic in her paper. Accordingly, she strongly believed that performance-based assessment should make use of simulated and real world applications and should be done in the nursing education. While concluding the paper, she highlighted that thanks to computer-assisted assessment technology, performance based assessment can be done for prospective nurses, and in this sense, their high-order thinking and problem-solving skills, referring to cognitive complexity, can be assessed in a desired way (p. 168).

Likewise these studies, Antoun, Nasr and Zgheib (2015) focused on the efficiency, applicability and comparison of paper-based testing, computer-based testing and automated response system with the participation of 45 3rd grade medicine students. All of the students attended three modules of clinical pharmacology course: firstly, they took traditional testing. Within the scope of the second and third modules, they took automated response system and computer-based testing. Then, in total, 40 students gave written feedback about the tools; accordingly, it was seen that the most interactive part was automated response system, and the students stated that computers could be used in order to support learning. Also, Pawasauskas, Matson and Youssef (2014) carried out a research with students and faculties in the pharmacy department at the University of Rhode Island that was about to implement computer-based tests within their curriculum. 66 students took computer-based tests in their elective courses, and before and after the test, they gave their preferences about this kind of test. At the end of the study, it was seen that a majority of the participants wanted to take courses that were integrated with computer-assisted testing systems.

Researchers also probed the relationship between prior computer knowledge and test takers' exam anxiety. Chin (1990) aimed to compare the participants' prior computer literacy and their computer-assisted test anxiety level as well as the relation between exam administration mode and their exam scores. 54 male and 51 female students at 10th grade took part in this study, and they took the multiple choice science exam either on computer or on paper. All students were given a questionnaire before the exam in order to define their computer anxiety level, and

after the exam the control group answered a questionnaire which was measuring their test anxiety level whilst the experimental group participated in a questionnaire which showed their prior computer literacy, computer anxiety and test anxiety. The results were the signs of the relation between prior computer knowledge and test anxiety, but computer test anxiety had no relation to previous computer knowledge, and groups showed no anxiety contradiction. The students' gender was not a prominent factor in the study. However, it's really surprising and different from other studies that the students who took computer-assisted tests became more successful.

Within a similar point of view, Kirsch *et al.* (1998) and Taylor *et al.* (1998) focused on the relation between test-takers' prior computer knowledge and their performances on the TOEFL CBT. After examining 90.000 test-takers who took the TOEFL CBT, it was concluded that more than 80% participants had at least moderate computer familiarity, and there was no significant effect of lack of computer knowledge over TOEFL CBT scores (as cited in Choi, Kim and Boo, 2003, p. 298).

Stricker, Wilder and Rock (2004) focused on the examinees' acceptance of the TOEFL CBT. This survey was not limited only to a country, but in total, 409 participants were chosen from three different countries; Brazil (Buenos Aires), Egypt (Cairo) and Germany (Frankfurt). These individuals took the computer-based TOEFL in the spring and summer in 1999. They were asked to participate in the questionnaire after completing their exams. The questionnaire consisted of different sub-scales such as TOEFL acceptance, acceptance of admissions tests, total computer attitude, total test anxiety, computer familiarity, preparation for TOEFL and admissions tests taken (pp. 39-40). At the end of the survey, it was deduced that in general, the test-takers in Brazil, Egypt and Germany showed positive interest in TOEFL CBT. This study indicated that computer-assisted assessment had started to appeal to people not only in USA but also in other parts of the world.

Dermo (2009) carried out a study about the students' perceptions on e-assessment. In this sense, 130 undergraduate students who took online tests during the academic year of 2007-2008 in England participated in the questionnaire. The research had six perspectives such as affective factors,

validity, practical issues, reliability, security and learning and teaching. According to the results, it was concluded that students had no anxiety towards the online assessment system; on the contrary, they had a desire to take online tests at university. Taking online tests was also regarded as a practical and secure domain by the participants. As for the validity and reliability of online tests, both factors seemed to be appropriate and related to the curriculum. Online tests, in addition, contributed to the participants a lot in terms of washback effect, clearly, the students were in thought of not only were they measured via online tests, but also they could learn with this smart tool. The researcher also aimed to see the relation between students' gender and attitudes as well as the relation between their age and attitudes. However, there was no statistically significant relation between students' age and their ratings, so it could be deduced that either the younger or the older participants they were, there would be no difference about their thoughts on online assessment. And, gender had no statistically significant relation, either.

Siozos *et al.* developed a computer-assisted assessment system called 'MyTest' with the contributions of both students and teachers with an aim to put forward a design for the challenges of CAA that were experienced in secondary schools in Greece (2009, p. 811). A total of 31 students and teachers participated for the design of the software system, and while doing so, they stated their reliance on 'MyTest'. Then, another group, made of 15 students and 16 teachers, was tested with this system. This group, again, found the system useful and reliable.

Teachers' readiness to use CAA in their classes was later examined by Laborda and Royo (2009) who were triggered with the Spanish Government initiative of passing computer-assisted university entrance examination (PAU). The researchers paid attention to the need of teachers' familiarity to the use of ICT in their classes and their training on that issue. Therefore, 26 teachers from different private schools took in-service teacher training course in Valencia. During that time, basic ICT strategies and the design of CAA were introduced to the teachers, and long discussion sessions were included during that training. At the end of the training session, the teachers were asked 10 questions so as to shed light upon the survey's aims, and it was found that although they were satisfied with the in-service training, they were in thought of not to be able to use CAA in their classes.

Laborda and Litzler (2011), in addition, used another survey to detect teachers' reactions to CAA. In this regard, 24 English language teachers, who were trained to use computer technologies as assessment tools, were included in the study. The aim was to define whether the teachers would change their teaching and assessment methodologies after the training session had been completed. Observation was utilized in this survey in addition to 10 items that were posed to the teachers on computer after the training. The results pointed out that organizing that kind of teacher training courses could be challenging, but definitely, it worked to increase teacher efficiency over implementing computer-based tasks, especially testing issues.

Within a different perspective, Al-Saleem and Ullah (2014) discovered the security features in computer-assisted assessment. They mentioned the gradually improving technology in testing tools and the importance of security. Therefore, they developed palm-based biometrics. Students are required to write their usernames and passwords; following to it, their palms are captured so as to allow them enter the system with security.

These are the summaries and compilation of the studies conducted abroad. Now, the related studies done in Turkey are given.

2.8.2. Turkey

Although there are not many related studies on computer assisted assessment or computer adaptive testing done in Turkey, some researchers have focused on the topic from different perspectives. All of them are grouped in relation to their research aims below.

Akdemir and Oğuz (2008) carried out a study with an aim to define any probable difference between students' scores in relation to test-mode: paper-based or computer-based exam. A total of 47 junior students at a state university in the Black Sea region participated in the study, and whilst the experimental group took the exam on computer, the control group had the traditional way of testing. At the end of the research, it was found that there was no significant difference between groups ($F = .153$, $p > .05$). Therefore, it can be deduced that computers can be alternatives as testing tools at Turkish higher institutions.

A similar study was done by Yaman and Çağıltay (2010). They conducted a study about the comparability of computer-assisted tests versus paper-based tests, specifically in engineering departments. In other words, the researchers aimed to answer whether students would perform better or worse once they were tested via computers, and with this aim, they worked with 209 engineering students in their 1st year at a private university in Ankara. Students were randomly assigned to two groups according to their last names, and whilst 96 students took exam on computer, 113 students did paper-based exam. The results showed that there was no significant difference between groups' scores depending on the exam-mode.

Lastly, Yağcı, Ekiz and Gelbal (2011) at a state university in Central Anatolia region carried out a similar study on this topic. This time participants were 75 vocational school students in the department of business administration. In order to reveal the probable academic success differences among participants, the exam was done in two ways (CAA versus PBT), and at the end, participants' scores were compared. It was found that students who had taken the computer-assisted exam succeeded 35% higher. This finding again contributes to the efficiency of computer-assisted assessment as well.

On the other hand, Semerci and Bektaş (2005) made a survey about how to improve the validity of web-based tests. First of all, the authors stated in their study that they strongly believed in the efficacy of web-based exams, and the related studies done abroad supported their belief. In this regard, they collected data from four different state universities (Anadolu, Sakarya, Fırat Universities and METU) in Turkey, where web-based testing was being administered. The researchers sent emails to a total of 45 people at those universities as to collect data for the study, and only 33 of them wrote back. After the data were analyzed, some ways to improve the validity of web-based tests were defined:

- ✓ Digital identities like fingerprint and voice control should be used.
- ✓ Teachers should encourage learners to make projects and research.
- ✓ Mini-quizzes and video-conferencing can foster learning, so teachers should make use of them in their courses.

Within a similar vein, Delen (2015) aimed to focus on how to increase the validity and reliability of computer-assisted assessment. In this sense, optimum item

response time for each question was shown on the screen when the participants were busy with answering the exam items, and the findings revealed that if students were offered optimum item response time, more valid and reliable tests would be achieved rather than paper-based tests.

Cömert (2008) and Erdogdu (2009) developed systems based on computer-adaptive testing principles. In his thesis, Cömert developed a software system, and among 176 9th-degree students in an Anatolian Technical high school in Istanbul, 20 of them were chosen randomly for the research. After those students had been assessed with CAT, it was concluded that time-saving was achieved since less questions were asked, comparing to the traditional way of testing. The reason for the decrease in the number of exam items were that the more each student succeeded during the exam, the less number of exam items were displayed on screen. Without any doubt, this bears a big advantage in education. Erdoğan (2009), also, used of CAT in her thesis research with an aim to define exam items' difficulties to be asked, because creating an item pool would improve the efficiency of computer-adaptive tests. In this vein, 368 students who took the midterm exam of *History of Civilization* course and 356 students who took the final exam of the same course constituted the sample of participants in this study. These two exams were administered as the computer-based version, and in relation to the responses given by students, item difficulties were ranged as *very easy, easy, middle, hard and very hard* as within an item pool that would be base for CAT.

Alkış (2010) made a survey about the factors affecting test-takers' attitudes towards computer assisted tests by using Technology Acceptance Model (TAM) in her study done with a total of 332 students. After taking the exam via computers, the students were asked to fill in a questionnaire, and it was found that perceived usefulness was the most motivating determinant in the study. Perceived ease of use had positive relation to perceived usefulness and intention, which displayed that if students took exam without challenges, they would intent to use it again and believe in the efficiency of it. In addition, once students had computer anxiety, their perception of ease, intention and perceived usefulness were affected negatively. In other words, when a student was anxious toward the technological tool, s/he would have no belief about its usefulness or its easiness to use, and so would not

have any desire to utilize it in the future. Lastly, it was emphasized that students' computer attitude and anxiety defined their attitudes towards computers as testing devices.

Oz (2014) also conducted a study with an aim to investigate prospective English language teachers' attitudes towards web-based testing. In this regard, in total, 50 students enrolled in a content knowledge course at a state university took web-based exam, and then, participated in a questionnaire. Moreover, interview was done since the participants could have a chance to ignite a different point of view related to the web-based exams. The results revealed that the student teachers had a positive attitude toward perceived ease of use and computer attitude, but were neutral with respect to their intent to take this sort of tests in future, which was supported by the low values given to perceived usefulness by the student teachers. Also, they wanted to take neither fully web-based nor fully paper-based exams, but a combination of these two exam modes would trigger their success and motivation for future classes. Although male participants stated that they spent much time on computer and had a high level of computer literacy, the results proved no statistical relation between the prospective English teachers' gender and their attitudes to web-based exams. Daily use of internet and the level of computer literacy were found to have a profound effect over the participants' thoughts about web-based exams.

Lastly, Akkoç (2012) touched upon a different point of view by focusing on the need of teacher preparation program for effective technology integration in classes. Therefore, the program, of which basis was *technological pedagogical content knowledge*, was developed, with great emphasis over *assessment via technology*. 41 prospective mathematics teachers at Marmara University took part in two different workshops during three months, and their active participation was needed. After workshops were done, participants were asked to share their opinions, and many of them stated their preference of computers as assessment tools, because fostering learning could be enhanced.

After reviewing all studies conducted both in Turkey and abroad, it is detected that there are a few related studies about the impact of computer-assisted assessment on students' achievement and attitudes. Therefore, it is a need to search and put forward the results concerning the efficiency of computers as testing tools and

test-takers' reactions, success and attitudes towards the system. In this regard, future applications and use of CAA will have a much more fertile ground.

2.9. Conclusion

Assessment is an indispensable part of education at all levels of education and directs the teaching process (Fisher and Frey, 2007). Therefore, a particular attention needs to be paid in order to administer valid, reliable and practical exams. In addition, washback effect, which either reveals the deficient points of test-takers or supports learning process, carries importance for teachers and students as well. Today, cutting-edge technology makes the administration of effective tests possible thanks to computer-assisted testing, web-based testing or computer-adaptive testing. Immediate feedback, objective scoring, time-control on screen, individualized assessment and grading items just upon the end of exam are given as the benefits of technology-based tests. Therefore, a variety of institutions around the globe have started to adopt computer systems as alternative assessment tools. However, some questions like students' scores comparison up to exam administration mode, equivalence of computer-based versus paper-based tests in terms of validity and reliability coefficients and what exam mode students prefer are arisen, in this sense, this study aims to search these points.

3. METHODOLOGY

3.1. Introduction

This chapter introduces the general research design of the present study and gives detailed information about the research model adopted. In addition, the chapter presents the research setting, participants and instruments involved in the study. Finally, the data collection procedure and data analyses of both quantitative and qualitative research methods are also presented in the related sections of the chapter.

3.2. Research Design

For about a century, most researchers have traditionally focused on either quantitative or qualitative research methods depending on the aims of their inquiries. Over a decade, however, researchers in social sciences have increasingly adopted a new research paradigm called the mixed-method research design, which integrates both quantitative and qualitative methods (Dörnyei, 2007). In other words, the data collection is supported by gathering both numeric information (e.g., on instruments) and textual information (e.g., on interviews) so that the researcher's final database includes both quantitative and qualitative information.

In quantitative research method statistical values are precedence while qualitative research method is based on non-numerical data. With respect to the purpose of study, either of the aforementioned research methods can be used. Each of these research methods has both advantages and caveats within their nature. As Dörnyei (2007) points out, “....quantitative inquiry is systematic, rigorous, focused and tightly controlled, involving precise measurement and producing reliable and replicable data that is generalizable to other contexts” (p. 34). On the other hand, qualitative research method offers some other kinds of advantages compared to the quantitative one. Qualitative research gives detailed descriptions of the study in contrast to fixed numerical, namely statistical data, focuses on an intensive study with fewer participants and tends to highlight new hypotheses or assumptions (Mackey and Gass, 2005). Similarly, Brown (2004, p. 485) also emphasizes that “one of the great strengths often cited for qualitative research is its potential for forming new hypotheses.” By integrating these two research

methods, the mixed-methods research contributes to a better understanding of research questions in the study. In other words, researchers who use this design compare or ‘triangulate’ the findings of both quantitative and qualitative research methods to see strong and weak sides of the study (Dörnyei, 2007, p. 43). Consequently, the researchers can combine the two types of research methods for complementary purposes.

The present study also employed the mixed methods research sequential design in order to define the convergence of data results. The first part included quantitative research methods. Accordingly, the participants were randomly assigned to experimental and control groups. While the experimental group received computer assisted test, the control group did not. After the administration of exam, participants in both groups responded to statements in a questionnaire of which results were calculated to obtain statistical data to search the research questions of the study. As seen, the first part covered the experimental research design without a pre-test. In some studies it may be ‘undesirable to give a pretest because the pretest itself might alert participants to what the treatment concerns’ (Mackey and Gass, 2005, p. 149). In this study, the researcher adopted the same assumption and used the experimental design-a posttest only. As summarized in Figure 3.1, in the second part of the study, a qualitative research method was utilized because it was the best way to describe some topics, mainly attitudes, in detail. This study aimed to investigate the attitudes of pre-service English language teachers towards a computer assisted assessment system. Obviously, an interview also made a great contribution to the study because it supported the statistical data analysis to some extent and gave a clearer understanding of technology use in assessment.

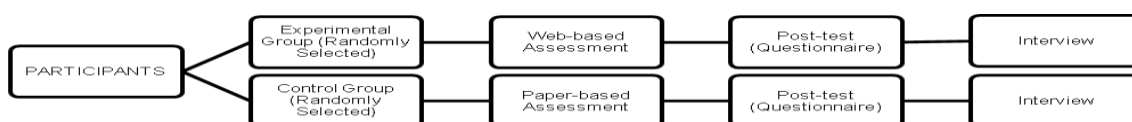


Figure 3.1. The Flowchart of the Research Model

3.2.1. Quantitative Research Design

The quantitative research model of the study covers the experimental study - a posttest only design. Accordingly, there is no place for pretests in the study, just

the posttests are used. The participants for the study were enrolled in *Approaches to ELT* in the department of English language teaching at Hacettepe University, and they took the content knowledge courses by the same instructor. After the researcher had randomly assigned the participants to two groups, the control group took the achievement test in a traditional way. On the other hand, the experimental group took the same exam through a computer-assisted system. When the exam was over, both groups were administered a questionnaire adapted from Alkış (2010) to state some background information of participants and their attitudes towards computer-assisted assessment. Alkış (2010) used the questionnaire only for those who took the exam by means of a computer-assisted assessment tool. However, the researcher aimed to detect the attitudes of the control group towards computer-assisted assessment system in this study, too. Hence, the questionnaire was adapted so as to state the control group's attitudes. Both versions of the questionnaire are presented in Appendix 2 and 3. Lastly, the validity and reliability values of the questionnaires are given in the following parts.

3.2.2. Qualitative Research Design

Numerous methods of data collection such as interviews, observations, ethnographies, and so forth are available for qualitative studies (Mackey and Gass, 2005). Of them, interview is of importance in this study as it gives an extended nature of the study from the viewpoint of each participant. Seidman (2006) holds that an in-depth understanding of other people's experience in a particular domain can be achieved thanks to interviewing.

In the present study, the interviews were conducted following the data collection. Two interviewees were invited in an office for each session. There were some questions prepared in advance by the researcher (see Appendix 4), and a semi-structured interview constituted the focal point with an aim to let participants explain their ideas without being limited to the scope of interview items.

3.3. Setting and Participants

The present study was conducted at Hacettepe University, one of the leading universities in Turkey, and a total of 97 students participated in the study. As stated before, the first part of the study made use of quantitative study which covered exam and questionnaire implementation. A total of 97 students took the

exam and responded to the statements in the questionnaire. As for the second part of the study, which was related to the qualitative-based research part, randomly assigned 26 test-takers were asked to take part in interview session. The tables below show the frequencies of the participants according to different variables such as gender, grade point average, daily internet use, computer literacy, and so forth.

As shown in Table 3.1., 77 females (79.40%) and 20 males (20.60%) took part in the study.

Table 3.1. Distribution of Participants by Gender

		<i>Frequency</i>	<i>Percent</i>	<i>Valid Percent</i>	<i>Cumulative Percent</i>
Valid	Female	77	79.4	79.4	79.4
	Male	20	20.6	20.6	100
	Total	97	100	100	

A total of 97 student teachers who were enrolled in *Approaches to ELT* course were randomly assigned to two groups. While doing so, simple random sampling was taken into consideration. As seen, a total of 50 (51.5%) test-takers took the exam on a computer-assisted assessment system (CAA) and 47 (48.5%) test-takers took it in a traditional, paper-based, way (Table 3.2).

Table 3.2. Distribution of Participants by Exam-mode

		<i>Frequency</i>	<i>Percent</i>	<i>Valid Percent</i>	<i>Cumulative Percent</i>
Valid	PBT	47	48.5	48.5	48.5
	CAA	50	51.5	51.5	100
	Total	97	100	100	

The researcher collected data about the participants' grade point averages by classifying them into three groups: students who got between 3.50 and 4.00; students who got between 3.00 and 3.49; students who got between 2.99 and below. Accordingly, the majority of the participants (N= 54, 55.7%) got between 3.00 and 3.49. On the other hand, the participants who got between 3.50 and above constituted the minority as just 3 students achieved it (3.1%). As for 40 participants, who were 41.2 percent of participants, they got 2.99 and below (Table 3.3).

Table 3.3 Distribution of Participants by GPA

		<i>Frequency</i>	<i>Percent</i>	<i>Valid Percent</i>	<i>Cumulative Percent</i>
<i>Valid</i>	3.50-4.00	3	3.1	3.1	3.1
	3.00-3.49	54	55.7	55.7	58.8
	2.99 and below	40	41.2	41.2	100
	Total	97	100	100	

The researcher was also motivated to investigate the participants' level of technology use and computer literacy. As seen in Table 3.4, a great majority of the participants (N= 50, 51.5%) use computer and internet more than 8 hours daily. The rest of them have close values in terms of daily use of internet as seen below.

Table 3.4. Distribution of Participants by Daily Use of Internet

		<i>Frequency</i>	<i>Percent</i>	<i>Valid Percent</i>	<i>Cumulative Percent</i>
<i>Valid</i>	Less than 3 hours	17	17.5	17.5	17.5
	3-6 hours	12	12.4	12.4	29.9
	6-8 hours	18	18.6	18.6	48.5
	More than 8 hours	50	51.5	51.5	100
	Total	97	100	100	

The present study collected data about all participants' computer literacy, and they were categorized as *bad*, *tolerable*, *good* and *very good*. The frequencies were respectively as follows: N= 6, 6.2%; N=50, 51.5%; N=34, 35.1%; N=7, 7.2%. As seen in Table 3.5, more than half of the participants saw their computer literacy as *tolerable*, and 34 participants had *good* level of computer literacy; that proved there were no obstacles to carry out the study since nearly all of the participants had enough computer literacy to take the exam via a computer-assisted assessment system.

Table 3.5. Distribution of Participants by Computer Literacy

		<i>Frequency</i>	<i>Percent</i>	<i>Valid Percent</i>	<i>Cumulative Percent</i>
<i>Valid</i>	<i>Bad</i>	6	6.2	6.2	6.2
	<i>Tolerable</i>	50	51.5	51.5	57.7
	<i>Good</i>	34	35.1	35.1	92.8
	<i>Very good</i>	7	7.2	7.2	100
	<i>Total</i>	97	100	100	

Furthermore, all participants took *Computer I* and *Computer II* courses in their first class which enabled them to acquire the necessary computer skills needed in the coming years. Consequently, the profile of participants' computer literacy supported that there was no need to prepare the participants by making them familiar with a computer-assisted assessment system, which required no great effort of the test-takers, before conducting the survey. However, the researcher prepared the necessary instructions so that the participants could better understand how to take the online achievement test and complete it without any challenges during the test-taking process. In addition, all the computers used for computer-assisted exam were checked in order to identify whether they had any technological problems or not, and the web-site by which the participants would access the exam-page was opened to make the process easy for the test-takers. Lastly, there were experienced proctors about the computer-assisted assessment system in each computer lab to help any test-taker in case it was needed. All in all, even though the test-takers' computer literacy and their prior experience displayed that they were accustomed to using computers at the desired level, the researcher took necessary measures so as to carry on the study under control.

Although participants' *computer literacy*, *daily use of internet* and *approximate time of starting to use computer* were described in general, independent samples t-test was run with an aim to compare participants in both groups in depth. As seen in Table 3.6 below, there were no statistically significant differences between groups. Firstly, participants in both groups did not differ in their level of computer literacy $t_{(95)} = -.098$, $p = .92 > .05$. Secondly, no significant difference was found in terms of participants' daily use of internet and/or computer $t_{(95)} = .359$, $p = .72 > .05$. Thirdly, the test-takers in both groups showed no significant difference as for their approximate time to start to use computer $t_{(95)} = .630$, $p = .53 > .05$. Consequently, it

can be easily mentioned that participants constitute a homogenous group. In other words, participants had similar features in both groups.

Table 3.6 Independent Samples T-Test

		<i>N</i>	<i>X</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>Sig. (2-tailed)</i>
<i>Computer Literacy</i>	<i>PBT</i>	47	3.426	.773	-.098	95	.922
	<i>WBT</i>	50	3.44	.675			
<i>Daily Use of Internet</i>	<i>PBT</i>	47	3.085	1.195	.359	95	.721
	<i>WBT</i>	50	3	1.143			
<i>Approximate Time of Starting to Use Internet/ Computer</i>	<i>PBT</i>	47	1.66	.668	.630	95	.530
	<i>WBT</i>	50	1.58	.575			

As for the qualitative study, a total of 26 test-takers were involved in the study as to specify their attitudes towards computer-assisted testing in depth. While choosing participants to interview, it was aimed to focus on the test-takers having taken the exam on computers as their thoughts and/or suggestions about it were of importance for both the present and future studies. Also, volunteer students were tried to be chosen for interview sessions. Apart from that, there were no criteria like gender, grade point average, computer literacy, and so forth to choose the participants for qualitative study. As shown in Table 3.7, 19 test-takers (15 females and 4 males) within the experimental group (CAA) and 7 test-takers (6 females and 1 male) within the control group (PBT) were interviewed.

Table 3.7 Distribution of Participants in Qualitative Study

	<i>Female</i>	<i>Male</i>
<i>Paper-based Testing</i>	6	1
<i>Computer Assisted Testing</i>	15	4

3.4. Data Collection Instruments

The present study employed three major types of research instruments to collect the required data. First, a course achievement test was administered to all students enrolled in the previously mentioned course (Approaches to ELT). Second, all participants were asked to complete a questionnaire so that they could indicate their perceptions of computer-assisted assessment. Third, a semi-structured interview was conducted to collect further data about their opinions on

computer assisted assessment systems. Each type of these instruments is respectively explained in the following sections.

3.4.1. Course Achievement Test

A 60-item achievement test was developed within the scope of sources used in the content-knowledge course by the instructor by focusing on *communicative language teaching*. Initially, an item pool of 70 questions in multiple-choice, gap-filling and true/false formats was generated in order to cover the major aspects of communicative language teaching. Then, the questions were revised by an experienced professor and a lecturer in the English language teaching department with the aim of supplying face validity of the test. As the experts suggested some revisions for questions, the necessary measures were taken and a total of 60 questions, which included 50 multiple choice items with four alternatives and 10 gap-filling items, were selected in order to pilot their reliability. In this regard, the pilot study was done with 9 prospective English teachers studying at the same university. Additionally, the test-takers in the piloting group had taken the same content-knowledge course *Approaches to ELT* before.

There are several ways of calculating the reliability coefficient for tests. Test-retest reliability, alternate or parallel forms and internal-consistency reliability are commonly used to calculate the consistency level of a test or instrument (Baykul, 2000; Büyüköztürk, 2011; Demircioğlu, 2007; Kan, 2013; Whiston, 2009). Since the achievement test under discussion was dichotomous in terms of scoring, the internal consistency level of the test was calculated by using the *split-half reliability method*, which supports that “instrument is given once and then split in half to determine the reliability” (Whiston, 2009, p. 53).

After the piloting group had completed the test, the items were divided in half as *the odd-numbers* and *the even-numbers* to minimize some probable problems that could be caused by fatigue or boredom of test-takers towards the end of the test or by a test becoming gradually difficult (Blerkom, 2009, p. 49; Ravid, 2011, p. 195; Whiston 2009, p. 54). Then, the reliability coefficient was calculated. The results of the test revealed that the achievement test had very good psychometric properties ($r=.90$) in terms of reliability (Table 3.8).

Table 3.8. Split-half Coefficient of the Achievement Test in the Pilot Study

Cronbach's Alpha	Part 1	Value	0.509
		N of Items	30
	Part 2	Value	0.708
		N of Items	30
Total N of Items			60
Correlation Between Forms			0.821
Spearman-Brown Coefficient	Equal Length		0.902
	Unequal Length		0.902
Guttman Split-Half Coefficient			0.896

In such reliability analyses, to clarify the point further, a correlation coefficient takes values between +1 and -1: the values close to 1 prove the test's (or instrument) reliability (Blerkom, 2009; Ravid, 2011; Whiston, 2009). Ideally, .90 or higher values are statistically very significant while interpreting the reliability coefficients. On the other hand, .60 or higher values can be acceptable in educational settings as teachers may have limited knowledge and time to prepare the tests (Blerkom, 2009, pp. 51-52). As for the split-half coefficient which gives the value belonging to the half of the test, and the Spearman-Brown coefficient which gives the reliability coefficient of the whole test, the coefficient values are .896 and .902 respectively; these values demonstrate the reliability of the test used in the present study.

Since the items prepared for the achievement test had a high level of correlation coefficient, the test was accepted as reliable, one of the milestones in developing a test. Then, both paper-based and computer-assisted versions of the exam were generated with the same questions that had been piloted before (see Appendix 5). In preparing the test on paper, the layout was of great importance. As for the computer-assisted assessment system, which offered different types of questions, was used to make the students take the achievement test online. On computer-assisted assessment system, 10 items were displayed at once. Figure 3.2 shows how items were displayed on the web:

HACETTEPE UNIVERSITY
To the leading edge... Toward being the best...

Achievement Test
Time left: 0:44:52

Question 1 of 60

Communicative Language Teaching aims to apply the theoretical perspective of the communicative competence by _____.

- ☐ A) acknowledging the independence of language and communication
- ☐ B) enabling learners to know the rules of language
- ☐ C) making communicative competence the goal of language teaching
- ☐ D) enhancing the memorization of fixed routines

[Clear selection](#)

Next

Figure 3.2. A Sample Item on the System

The test-takers had the chance to move back and forth or skip any question as on paper-based version and to control the allotted time on their screen as seen in Figure 3.2. As for the gap-filling part, the test-takers were supposed to write the correct or possible answer in the given blank. All mandatory or optional answers were mentioned while editing the items on the web (Figure 3.3). Capitalization or punctuation mistakes were not taken into account for scoring.

[Edit question](#) [Preview question](#)

Question

The _____ version of Communicative Language Teaching claims that language is acquired through communication, so that it is not merely a question of activating an existing but inert knowledge of language, but of simulating the development of the language system itself.

Accepted answers

Add each separate accepted answer per box

Users will not see these when answering this question.

Mandatory

Optional

Optional

Optional

Optional

Figure 3.3. A Sample Item on the System

The participants who took the online version of the test logged into the system by entering their student IDs and passwords that had been previously prepared by the researcher (Figure 3.4). Once they logged in, they were required to write their

full names and email addresses. Before starting the exam, brief information about the exam (the allotted time, the number of questions, the cut-off point for passing the exam, whether they could skip the items or not) was displayed on the screen.

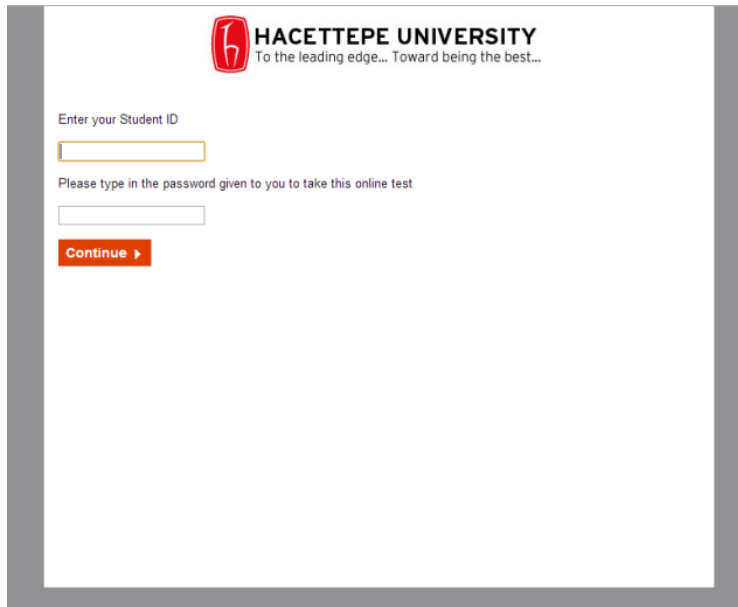
The image shows a web-based login interface for Hacettepe University. At the top, the university's logo (a red shield with a white 'H') is next to the text 'HACETTEPE UNIVERSITY' and the tagline 'To the leading edge... Toward being the best...'. Below this, there is a section for user authentication. It starts with the prompt 'Enter your Student ID' followed by a text input field. Below that is the prompt 'Please type in the password given to you to take this online test' followed by another text input field. At the bottom of this section is a red button with the text 'Continue' and a right-pointing arrow.

Figure 3.4. Log-in Screen

The computer-assisted assessment system offered several benefits both for the test-takers and the teachers: firstly, the test-takers could see their scores on screen just upon completing the exam. In addition, they could get immediate feedback, namely saw correct answers, for each item after confirming the questions. Secondly, once they achieved 70 or higher correct answers, a certification appeared on screen so as to motivate them. Lastly, the system stored all test-takers' responses, and the researcher reached them whenever needed.

3.4.2. Questionnaire

It is an undeniable fact that technology can be integrated within educational curriculum in relation to the learners' acceptance of technology. In this respect, for decades, a wide range of theoretical models has been proposed, "including innovation diffusion theory, the theory of reasoned action, the theory of planned behavior, and technology acceptance model (TAM)" so that the level of technology-users' acceptance can be defined (Saade and Bahli, 2005, p. 317). *Technology Acceptance Model* by Davis (1989), based on the *Theory of Reasoned Action* and focused on the perceived usefulness, perceived ease of use and attitudes (Saade and Bahli, 2005, p. 318), is the most accepted one among

them. For this reason, the instrument used in the present study utilized this aforementioned model, TAM.

Two of the main components of TAM are *perceived ease of use* and *perceived usefulness*, and Davis (1989, p. 320) defines the former as “*the degree to which a person believes that using a particular system would be free of effort*”, and the latter as “*the degree to which a person believes that using a particular system would enhance his or her job performance*”. Additionally, *attitudes* are of importance in TAM as mentioned in Figure 3.5 below. Consequently, it is quite clear that there are some dimensions that make people accept or reject the state-of-art technology, namely use of technology for various purposes (Davis, 1989, p. 320). Swanson (1987) claims that items like “*important*”, “*relevant*”, “*useful*”, and “*valuable*” indicate a tendency towards perceived usefulness while items like “*convenient*”, “*controllable*”, and “*easy*” refer to perceived ease of use (cited in Davis, 1989, p. 322). The relation between these two components can be depicted below (Park, 2009, p. 150).

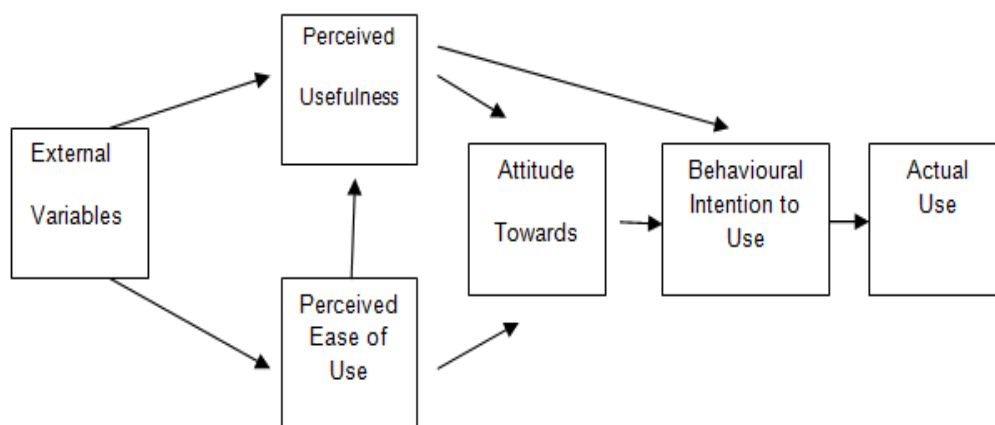


Figure 3.5. Technology Acceptance Model

By taking into account all studies previously done, Davis (1989) developed a scale with 14-item and tested it. After the statistical data for reliability and validity was collected, 6-item for perceived ease of use and 6-item for perceived usefulness were put forward. Further studies were conducted in order to analyze the model because of its wide acceptance in information technologies-related studies. Park (2009) confirmed that TAM helped researchers define to a large extent the attitudes of people towards computer-assisted studies or tasks, and that perceived usefulness had a direct effect over *intention*, which was a part of the present

study's instrument, to utilize technology-based tasks in contrast to the original model, TAM. Ong and Lai (2006) conducted an extended survey based on TAM by focusing on probable gender difference of users' attitudes towards or acceptance of technology. Computer self-efficacy (CSE), perceived ease of use (PEU), perceived usefulness (PU) and intention to use (INT) were the sub-scales of the instrument in their study. Accordingly, CSE, referring to a user's belief in successfully performing a computer-based task, was highly in relation to perceived usefulness and had a significant effect over perceived ease of use (Ong and Lai, 2006, pp. 818-819). As for the gender difference by CSE, women commonly showed high level of *computer anxiety*; clearly, women demonstrated a lack of confidence or self-efficacy in performing or using computers. Hence, it can be mentioned that there is a reversing relation between self-efficacy and anxiety. Lee *et al.* (2005, p. 1095) carried out a survey that was based on TAM and integrated both extrinsic (perceived ease of use and perceived usefulness) and intrinsic (perceived enjoyment) motivators to determine the level of users' *intention* to use technology. The results of their study proved that there was a strong and direct effect of perceived ease of use and perceived usefulness over intention to use technology.

Alkış (2010) adapted the survey items of various studies based on *technology acceptance model* and generated a survey instrument so as to identify the technology acceptance level of the students at a large state university in Turkey. The questionnaire consisted of 37 items and five sub-scales: *perceived ease of use*, *perceives usefulness*, *intention*, *anxiety* and *computer attitude*. The participants chose the best option ranging from (1) *Strongly disagree*, (2) *Disagree*, (3) *Neutral*, (4) *Agree*, and (5) *Strongly agree*. In her study, Cronbach's alpha coefficient was found to be .91, proved that the internal consistency of the instrument was very high.

The present study used the adapted and validated survey instrument by Alkış (2010) for the experimental group that took the computer-assisted exam. On the other hand, the researcher adapted the questionnaire for the control group that took the exam in traditional way by asking them to suppose they would take the exam on computer as to collect the data from this group as well. The

questionnaires can be found in Appendices A and B. The related data results are given below.

In some studies, like five-point Likert shaped questionnaires, the items are not graded or scored dichotomously. In case of it, Cronbach's alpha is the best formula to calculate reliability coefficient (Blerkom, 2009, p. 50). While interpreting the calculated value, .70 or higher coefficients are the indicators of a reliable instrument; however, .60 or higher values are acceptable in educational setting as well (Blerkom, 2009, pp. 51-52). In this sense, reliability coefficient of survey instrument, done with the experimental group, is given in Table 3.9. As seen, overall items have a high level of alpha coefficient .957 that is so close to +1 that it proves the internal consistency of the instrument. When the reliability coefficients both of each-subscale and of each item were analyzed, it could be stated that the instrument was reliable and measured what it was intended to measure, in other words, the attitudes of test-takers taking the exam on a computer assisted system towards computer-assisted testing were measured with a reliable instrument.

Table 3.9. Reliability Coefficients of Computer-assisted Assessment Questionnaire and Its Factors

<i>Factors</i>	<i>Number of Items</i>	<i>R</i>
<i>Factor 1 - Intention</i>	11	.963
<i>Factor 2 - Perceived Usefulness</i>	5	.920
<i>Factor 3 - Perceived Ease of Use</i>	12	.812
<i>Factor 4 - Anxiety</i>	4	.771
<i>Factor 5 - Computer Attitude</i>	5	.859
<i>Whole Instrument</i>	37	.957

The original instrument generated by Alkış (2010) was about the attitudes of the participants who were taking or had taken any computer or web-based tasks. However, the present study also aimed to define the attitudes of the control group towards computer-assisted assessment systems in case they could use technology-related tasks, namely computer-assisted exams. Because of this reason, the researcher reworded the instrument and used it with the control group in order to compare the results of the instruments belonging to both experimental and control groups so as to clarify whether any statistically significant difference between two groups occurred or not. Overall, the instrument indicated a high level of reliability with an alpha coefficient .943. Similarly, the factors had high a level of coefficients.

Table 3.10 Reliability Coefficients of the Adapted Version of Computer-assisted Assessment Questionnaire and Its Factors

<i>Factors</i>	<i>Number of Items</i>	<i>R</i>
<i>Factor 1 - Intention</i>	11	.939
<i>Factor 2 - Perceived Usefulness</i>	5	.842
<i>Factor 3 - Perceived Ease of Use</i>	12	.875
<i>Factor 4 - Anxiety</i>	4	.596
<i>Factor 5 - Computer Attitude</i>	5	.824
<i>Whole Instrument</i>	37	.943

In order to analyze the relation between the two instruments, the *Pearson Correlation Coefficient* was calculated and a high level of coefficient was found ($r = .821$) that was the proof of the fact that the original questionnaire and the adapted version of the questionnaire mutually reinforced each other (Table 3.11).

Table 3.11 Pearson Correlation Coefficients of the Questionnaires

	<i>Number of Items</i>	<i>R</i>
<i>Computer-assisted Testing Questionnaire</i>	37	0.821
<i>Adapted Version of Computer-assisted Testing Questionnaire</i>		

In summary, from the reliability analyses conducted, both the questionnaire applied to the experimental group and the adapted questionnaire administered to the control group indicated that the two instruments demonstrated good psychometric properties for the research purposes.

3.4.3. Interview

Two major types of research models, quantitative and qualitative, may require the researcher to employ different kinds of data collection and analysis techniques. In other words, the aims of the research and the instruments to be used undoubtedly play a significant role in the selection of research methodology. In recent years, researchers have started to use mixed-methods research “to triangulate both more quantitative and more qualitatively oriented data” (Mackey and Gass, 2005 p. 307). In this regard, employing mixed-methods research gives clear and supported results. In addition, some study topics, especially the ones about attitudes as is the case in the present study, may be limited by the use of only quantitative data. As Mackey and Gass (2005, p. 307) hold, “when included in a primarily quantitative

report, qualitative data or analytic techniques may provide unique insights ... Similarly, qualitative reports may become clearer when some quantitative analysis is included". The related literature about research methods and the topic area of the study guided the researcher to utilize mixed-methods research. In this respect, after the test-takers in both groups had taken the exam either on paper or on computer, they were asked to answer the related questionnaire which constituted the base for quantitative data collection. As the last phase of gathering data, interview, a kind of qualitative data collection techniques, was done with convenience sampling that referred to the test-takers in both groups, who were willing to participate in an interview session, in order to collect a large-set of data. Hence, the researcher collected the related data as to support the numerical data results and as to state a deep-insight into the attitudes of test-takers.

Given the overall importance of the triangulation in the research, the researcher prepared a semi-structured interview that contained a total of seven questions. However, the participants who engaged in the interview were also given the chance to touch upon the pros and cons of computer-assisted assessment systems even if they were not mentioned within survey items as the semi-structured interview was modeled. The related questions or items are given below:

Interview Questions

- 1. Have you ever taken any formal computer-assisted exam before? If so, what was/ were your reaction(s) towards it?*
- 2. What did you like and/or dislike about the computer-assisted assessment?*
- 3. If you have to make a choice over a computer-assisted exam and a traditional paper and pencil exam, which one would you like to take? Why?*
- 4. What advantages do you think a computer-assisted exam has over a paper-based test?*
- 5. What disadvantages do you think a computer-assisted exam has over a paper-based test?*
- 6. Does using a computer-assisted system as an assessment tool save time?*
- 7. Do you think the computer-assisted assessment (specifically web tools) should replace the traditional paper-and-pencil assessment? Why or why not?*

As seen above, the interview items or questions were prepared in relation to the survey items for giving support to the statistical results. Also, each interviewee could have the chance to express his/her ideas about the computer-assisted assessment system without any limitations. For instance, every student has different learning styles, and undoubtedly this may affect their ways of taking exam: while some students need to underline or highlight some words or sentences on a paper, some students do not need to do it. During the interviews, the interviewees could touch upon all these points in depth and give suggestions for further studies. Consequently, the researcher collected data by referring to the different points of views of test-takers not only for supporting the data results, but also for giving suggestions or planning for further studies.

3.5. Data Collection Procedure and Analysis

This part is devoted to giving a detailed picture of data collection procedure and a general view of phases in data analyses.

3.5.1. Data Collection Procedure

Although the present study was based on the relation between two different exam modes, the data collection procedure included the content-knowledge course and the achievement test (done both on computer and on paper). Later on, both quantitative and qualitative data collection instruments were employed to adequately gather information.

Content-knowledge Course

Prospective teachers of English in Turkish higher education system need to take a number of pedagogical content knowledge courses. As part of the pedagogical knowledge base for English teacher trainees, the department also offers a course that specifically focuses on language teaching approaches, methods and techniques. This course is Approaches to ELT. The general objectives of this course can be stated as follows:

Enabling students to have an understanding of such basic approaches, methods and techniques as Communicative Language Teaching, Multiple Intelligences Theory, Cooperative Language Learning, Content-Based Instruction / Content and Language Integrated Learning, Task-Based Instruction, Learning Strategy Training, and demonstrating students how they would apply them in real teaching contexts and integrating technology into the language classroom.

Regarding the learning outcomes for the course, at the end of the course, the students should be able to identify and describe such major language teaching methods such as Communicative Language Teaching, Multiple Intelligences Theory, Content-Based Instruction / Content and Language Integrated Learning, Content – Based Instruction, Task-Based Instruction, Learning Strategy Training, distinguish the techniques that are used in language teaching methods, demonstrate the use of a specific technique in a particular language teaching method or approach, and identify ways of using technology appropriate to their instructional settings.

During the data collection phase, this course was offered in three different sections, taught by the same faculty member and taken by a total of 97 students in the department. Prior to the administration of the achievement tests and questionnaires, the course instructor devoted a total of nine hours in three weeks to communicative language teaching (CLT) as the first module in the course. In the first week, he lectured in multiple sections of the course on the development of CLT, its basic principles and assumptions, and a variety of techniques that are commonly employed in today's communicative language teaching. There was also a video demonstrating a typical use of CLT in an ESL classroom. For the purpose of performance-based assessment, he had previously assigned some presentation tasks to a group of voluntary students. Therefore, in the second week, a group of student teachers briefly demonstrated how to conduct a language lesson based on some principles of CLT. Group presentation was further supported by a review of CLT principles and a question-and-answer session focusing on the observations attained from the demonstration. Finally, in the last three hours of the module the class sessions contained follow-up activities in which CLT was further reviewed through an evaluation of its basic principles and assumptions, teaching techniques, and some discussions questions.

In conducting the aforementioned activities, the students and the instructor basically used Diane Larsen-Freeman and Marti Anderson's (2011) *Techniques and Principles in Language Teaching*, an updated third edition of the popular introduction to language teaching methodology, which describes different methods and approaches in language teaching. Before student teachers engaged in their own group presentation, the course instructor had supported the CLT course

module by demonstrating a video originally developed for the first edition of the book in a joint project by Diane Larsen-Freeman and the U.S. Information Agency in 1990. The 20-minute video gave student teachers a general understanding of a CLT-based language classroom in action, but it was a bit outdated and lacked many aspects of today's CLT. Therefore, the course instructor also utilized content from Jack C. Richards' (2006) updated version of *Communicative Language Teaching Today* and Jack C. Richards and Theodore S. Rodgers' (2001) *Approaches and Methods in Language Teaching*.

The Achievement Test and Questionnaires

After all instructional activities were completed in the regular teaching sessions, the students were supposed to get the required information about CLT, and they needed to take an achievement test as part of the assessment process. Then, an item pool was planned to be generated so as to choose the reliable and valid items within it to ensure the exam's reliability and validity. The preparation of the test questions included as many different types of questions as possible such as multiple-choice, true/false, gap-filling, and so forth. Therefore, after the item pool, which contained a wide range of types, had been created, the questions were revised and piloted. According to the results of piloting, revision was needed for some questions, and some of items had to be omitted. The exam questions can be found in Appendix 5.

As mentioned previously, the participants were randomly assigned to experimental and control groups. Afterwards, they were all informed about when the exam would be taken and where they would take the exam since the experimental group would take it in three different computer labs while the control group would take it in two different rooms. In order to minimize or prevent any technical problem that was likely to occur during the test administration, all the computers had been checked before. In addition, there was an experienced teacher in each computer lab.

In the control group, the questionnaires were given to the participants as soon as they completed answering the questions. Similarly, each test-taker could access the questionnaire prepared online after confirming the exam questions in the experimental group. Both the original and adapted versions of questionnaire were

in English. Actually, it was assumed that the participants would have no difficulty in getting the meaning for each item as they were prospective English language teachers; however, in case there was any vocabulary they did not know, proctors explained the item so as to make them give trustworthy answers.

Interview

The interview for the study was the last phase of data collection procedure. As the interview would take a long time to conduct, it was fairly logical to include a few participants as possible. A total of 26 test-takers from both groups participated in the interview session. Clearly, students' GPA, gender, exam score, and so forth were not taken into account. On the other hand, it should be mentioned that there was a crucial criterion: volunteer participants had to be chosen. If not, the researcher could not have gathered enough data to support the numerical results and to give suggestions for further studies as well.

The interviews were carried out in an office in a relaxed atmosphere for the participants. Each interview was conducted in Turkish and lasted for approximately 20 minutes. During interviews, initially the researcher gave brief information about the process and let the interview look through the interview template (see Appendix 4). The interviewees started to speak when they were ready. The researcher directed the interview by focusing on the semi-structured items. However, there was not a strict rule of being stick to the questions; rather the participants were allowed to add different comments from their viewpoints. Voice-recording was done since it was so difficult for the researcher to take down and to guide the interview.

3.5.2. Data Analysis

The present study utilized a mixed-methods research model. Therefore, both quantitative and qualitative data analyses were employed to answer the research questions of the study. There were two instruments used for quantitative data collection: the questionnaire and the achievement test. In order to analyze the gathered quantitative data, they were entered into Statistical Package for Social Sciences (SPSS). While coding the questionnaires in SPSS, the researcher paid attention to reverse-coding the negative statements so as to reveal trustworthy results.

Before analyzing the gathered data on SPSS, it is a must to measure whether the data are normally distributed or not so as to find and use the most suitable statistics for the research questions, and there are many ways to do it: Skewness and Kurtosis, Levene test, Kolmogorov-Smirnov test (K-S test) and Shapiro-Wilk test are commonly used in order to reveal the data distribution. Among them, K-S test is generally used in case the study group includes more than 50 participants (Büyüköztürk, 2011, p. 42).

As this study had 97 participants in total, K-S test was run in order to define the normality level of the instruments, the questionnaire and the achievement test, with an aim to define what kinds of test would be used for analyzing the data. The table and figures below show the normality level of the questionnaire. Accordingly, it is seen that data are normally distributed and parametric statistical tests can be used so as to analyze research questions for significant value ($p = .200$) is higher than .05 ($p = .200 > .05$).

Table 3.12 Normality Test of the Questionnaire

	<i>Kolmogorov-Smirnov</i>			<i>Shapiro-Wilk</i>		
	<i>Statistic</i>	<i>df</i>	<i>Sig.</i>	<i>Statistic</i>	<i>df</i>	<i>Sig.</i>
<i>Instruments</i>	0.049	97	.200 [*]	0.989	97	0.598

^{*} $p > .05$

In addition to the numerical data given above, histograms and normal Q-Q plot offer visual cues about the test of normality, and the figures below show the normality test belonging to the questionnaire and support the numerical result shown above.

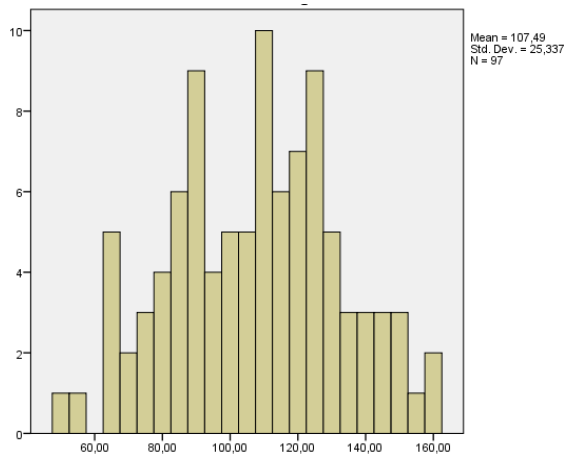


Figure 3.6. Histogram for Normality Test of the Questionnaire

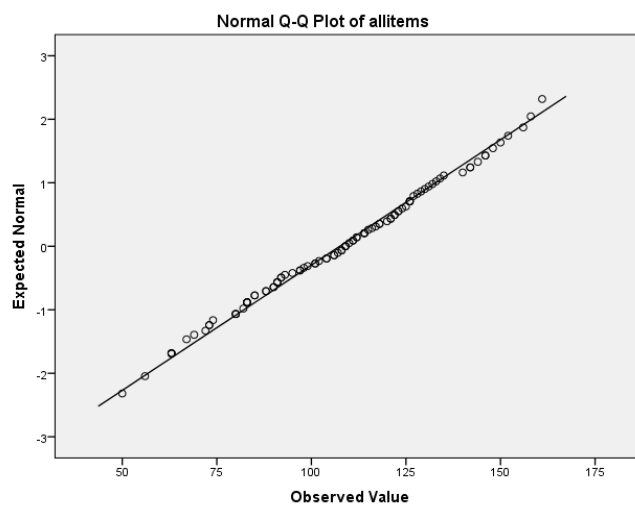


Figure 3.7. Normal Q-Q Plot for Normality Test of the Questionnaire

This study also made use of the achievement test, and there were some research questions on it. Therefore, measuring its normality level was needed as well. K-S test results given below display that the data are not normality distributed ($p = .002 < .05$), so non-parametric tests are done in order to find answers to the related research questions.

Table 3.13 Normality Test of the Achievement Test

	<i>Kolmogorov-Smirnov</i>			<i>Shapiro-Wilk</i>		
	<i>Statistic</i>	<i>df</i>	<i>Sig.</i>	<i>Statistic</i>	<i>df</i>	<i>Sig.</i>
Instruments	0.117	97	.002*	0.968	97	0.020

* $p > .05$

Also, histograms and normal Q-Q plot given below (Figures 3.8 and 3.9) offer visual cues about the test of normality and prove that the tests are not normally distributed.

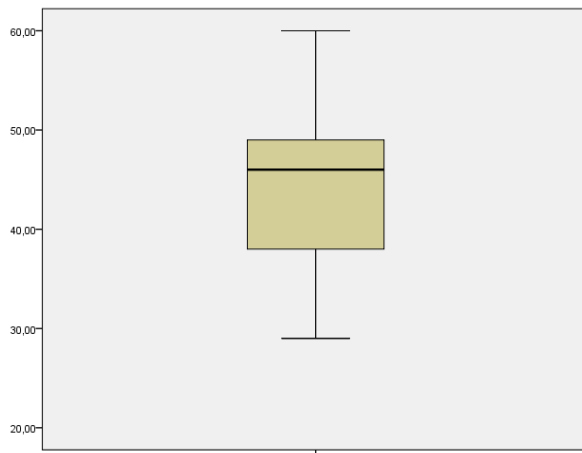


Figure 3.8. Histogram for Normality Test of the Achievement Test

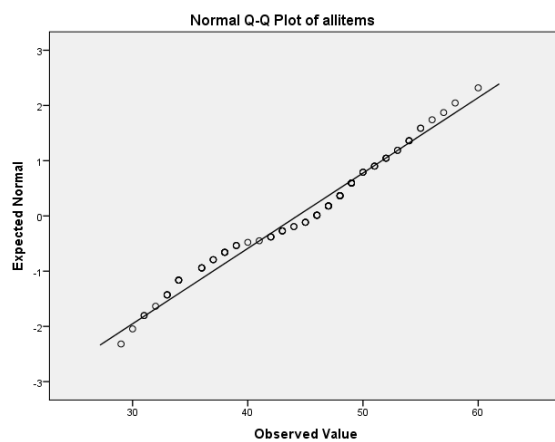


Figure 3.9. Normal Q-Q Plot for Normality Test of the Achievement Test

Initially, the descriptive statistics, including percentages, frequencies and mean scores of participants under a variety of criteria were calculated. Then, the reliability coefficient of test items was determined via Spearman-Brown formula. As there were two different versions of the questionnaire, namely two different instruments, the reliability coefficient of each instrument was needed; in this regard, Cronbach-alpha coefficient was calculated. In addition, Pearson-product Correlation was calculated in order to see the correlation between two questionnaires.

As for the data analyses of the research questions, the findings of independent samples t-test were given to reveal whether there was any statistically significant difference between the prospective English teachers' general attitudes towards computer-assisted assessment for their content knowledge course. In addition, Pearson Correlation was calculated so as to find the relation between the prospective English teachers' attitudes towards CAA and their gender, GPA, daily

use of computer and computer literacy. The probable difference of the test-takers' preference on exam mode was revealed with Chi-Square Test. Moreover, Mann-Whitney U test was done, because showing whether the test-takers' scores differed according to their exam mode or not was aimed. Lastly, split half reliability method and Spearman's rank order were calculated to determine the reliability and validity of the exams administered in two versions.

The last research question, as mentioned before, was about the general attitudes and opinions of the participants towards the computer-assisted and paper-based tests. Interview, one of the qualitative data collection techniques, was done in a semi-structured way so that each interviewee could explain the effect of the study over themselves. Although the interview items were written in English, the researcher tended to conduct the interview in Turkish, the mother tongue of the participants, so as to get them to explain their ideas without language anxiety. After each interviewee had been allowed to look through the items, voice-recording was made, and the gathered data were transcribed and translated into English.

Qualitative data analysis part was done by touching upon all interview items one by one. While doing so, the grounded theory principles were taken into consideration (Corbin and Strauss, 1990); the interviewees' responses or comments were categorized according to their similarities and differences. Then, selective coding was administered. It is one of the coding methods under the ground theory and focuces on "the main analytic idea presented in the research" (Corbing and Strauss, 1990, p. 14). Also, the qualitative data analysis part started when the researcher made interview with the first student, because as stated before, the semi-structured way was obtained in order to collect qualitative data. Therefore, the students and the researcher were not strictly tied to the interview items, but the researcher guided or posed some different questions that were raised from the comments given by the interviewees.

3.6. Conclusion

This part explains the research design, data collection tools and data analysis. Briefly, it used mixed methods research design, combination of quantitative and qualitative research methods. In order to collect data, prospective English

language teachers were randomly assigned to two groups, experimental versus control groups, and the students in the experimental group experienced the computer-assisted assessment whilst the rest students did not. After the exam, all participants were wanted to rate a questionnaire which had items about their attitudes toward CAA. Furthermore, some students in the experimental group were posed questions within the frame of interview. Following to it, the next chapter offers the related data analysis.



4. RESULTS OF DATA ANALYSIS

4.1. Introduction

This chapter presents the related data analyses under two broad headings referring to quantitative and qualitative data results. Firstly, inferential data analyses related to the questionnaires and the achievement test administered to two different groups were given. Secondly, data gathered from interviewees were analyzed in relation to the interview items.

4.2. Quantitative Data Analysis

Quantitative data collection tools included an achievement test and two different questionnaires, based on TAM (Technology Acceptance Model), regarding any difference in terms of test-takers' attitudes towards computer-assisted assessment systems. Both questionnaires had five factors such as *perceived ease of use*, *perceived usefulness*, *intention*, *computer attitude* and *anxiety*. The inferential statistics of survey items were shown by referring to the research questions as the study aimed to define whether there was any statistical significant difference between the experimental group and the control group in terms of these aforementioned five factors or not. In addition, the test-takers' performance in relation to the exam modes and their success were analyzed. Lastly, whether both exam versions overlapped in terms of validity and reliability or not was defined.

4.2.1. Results of Data Analysis for Research Question 1

This study aimed to find out the prospective English teachers' attitudes towards computers as testing tools. While doing so, all students participated in a questionnaire after they had completed their exams. This questionnaire had five factors, each of which focused on a different perspective.

What are the prospective English teachers' attitudes towards computer-assisted assessment for their content knowledge course in terms of intention?

The first factor of the questionnaire was about the test-takers' intention. Accordingly, the participants' thoughts, experiences and willingness on CAA were explained. The independent samples t-test was run and it was seen that there was

a statistically significant difference between the experimental and control groups $t_{(95)} = -3.36$, $p = .001 < .05$ as p value was lower than .05.

Table 4.1. Independent Samples T-Test for Factor 1 – Intention

		<i>N</i>	<i>X</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>Sig. (2-tailed)</i>
<i>Intention</i>	<i>PBT</i>	47	26.11	8.29	-3.36	95	.001
	<i>WBT</i>	50	32.94	11.36			

As seen in Table 4.1, the experimental group ($X = 32.94$) showed higher tendency and was much motivated towards taking tests on computers rather than the control group ($X = 26.11$). Accordingly, the participants in the experimental group showed higher tendency towards the efficiency of computer assisted assessment system and they liked the idea of using computers as testing tools. In general, the experimental group had a good experience on having been tested on computers. The students who did not have the chance to experience computers as assessment tools were not generally positive towards the system. Therefore, it is quite explicit that they did not show tendency to take computer assisted assessment for future courses.

What are the prospective English teachers' attitudes towards computer-assisted assessment for their content knowledge course in terms of perceived usefulness?

The second factor of the questionnaire is about the test-takers' perceived usefulness. In other words, the related statements in the factor 2 give values about the effectiveness of the computer-assisted assessment system, whether it helped the students increase their performance during the exam and/ or what their motivation on taking this kind of exam before and for future courses is. The table 4.2 below gives the related findings.

Table 4.2. Independent Samples T-Test for Factor 2 – Perceived Usefulness

		<i>N</i>	<i>X</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>Sig. (2-tailed)</i>
<i>Perceived Usefulness</i>	<i>PBT</i>	47	10.19	3.35	-2.86	95	.005
	<i>WBT</i>	50	12.66	4.93			

Table 4.2 shows the findings. As seen, there was a significant difference between groups $t_{(95)} = -2.86$, $p = .005 < .05$. Clearly, the control group had no belief on the usefulness of the computer-assisted assessment system and they had no desire to take computer based tests in the future. However, the experimental group found the system useful and showed desire to take that kind of testing systems for future courses.

What are the prospective English teachers' attitudes towards computer-assisted assessment for their content knowledge course in terms of perceived ease of use?

As computer assisted tests are not utilized commonly in Turkey, many of the students are not accustomed to using them. The related statements to this factor, perceived ease of use, give the values belonging to the participants about the easiness of using computer-assisted testing system. The experimental group used the system, and they stated their opinions whether it was easy or hard. On the other hand, the control group did not experience the computer-assisted testing system, but they participated in the questionnaire by assuming that they had taken or took this sort of test. In general, as seen in Table 4.3, there was a statistically significant difference between groups $t_{(95)} = -6.21$, $p = .00 < .05$.

Table 4.3. Independent Samples T-Test for Factor 3 – Perceived Ease of Use

		<i>N</i>	<i>X</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>Sig. (2-tailed)</i>
<i>Perceived Ease of Use</i>	<i>PBT</i>	47	30.08	7.33	-6.21	95	.00
	<i>WBT</i>	50	39.72	7.90			

It can be deduced that the experimental group ($X = 39.72$) was satisfied with the system whilst the control group ($X = 30.08$) still believed in the easiness and usefulness of the paper-based assessment. Therefore, many participants in the experimental group could use the system without any difficulty. More specifically, the students having been tested with the computer assisted assessment system found it easy to use and not so much effort was needed to use it.

What are the prospective English teachers' attitudes towards computer-assisted assessment for their content knowledge course in terms of anxiety?

The computer assisted assessment systems are not so common in Turkey that virtually all participants in the experimental group were tested on this system for the first time in this study. However, the control group still did not experience this smart testing system; therefore, the students may be anxious about the computer-assisted tests. In this regard, there were some statements for giving values about all participants' anxiety levels in both groups. As these statements had negative meanings, the researcher paid attention to reverse coding.

Table 4.4. Independent Samples T-Test for Factor 4 – Anxiety

		<i>N</i>	<i>X</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>Sig. (2-tailed)</i>
<i>Anxiety</i>	<i>PBT</i>	47	10.36	2.79	-2.44	95	.17
	<i>WBT</i>	50	12.00	3.72			

As shown in Table 4.4, there was no statistically significant difference between groups in terms of their anxiety to use CAA as significant value is higher than .05, $t_{(95)} = -2.44$, $p = .17 < .05$. In this regard, both groups might have fear of intimidating side of computer assisted assessment systems. This value may be the result of the fact that the students in the experimental group saw and used the system for the first time in this study. However, the control group still had no experience about being tested on computers.

In general, the control group tended to be anxious about the computer assisted assessment system. The reason for it might be that they participated in the questionnaire without having seen the system or been tested via it. Though the experimental group used the computers as testing tools, many of them had no prior experience, that's why they showed anxiety towards the system.

What are the prospective English teachers' attitudes towards computer-assisted assessment for their content knowledge course in terms of computer attitude?

The last factor of the questionnaire is about computer attitude. The students' thoughts on computers as smart tools and their benefits to our lives were investigated by touching upon the values of the related statements. In general, it can be seen in Table 4.5 that both groups believe the benefits of computers and

their positive impact on our lives and there was not any statistically significant difference between them $t_{(95)} = -.98$, $p = .330 < .05$.

Table 4.5. Independent Samples T-Test for Factor 5 – Computer Attitude

		<i>N</i>	<i>X</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>Sig. (2-tailed)</i>
Computer Attitude	PBT	47	18.42	3.32	-.98	95	.330
	WBT	50	19.16	4.00			

When the result was analyzed in depth, it could be mentioned that the students who were tested on computers and the students who were not thought that computers contributed a lot to our lives. Within a similar perspective, both groups agreed with the increase in the life standards thanks to computers and with the benefits of continuously growing opportunities of computers.

4.2.2. Results of Data Analysis for Research Question 2

What is the relation between prospective English teachers' attitudes of computer-assisted assessment and their demographic factors such as gender, GPA, daily use of computer and computer literacy?

After giving the prospective English teachers' attitudes toward the computer assisted assessment in relation to five dimensions within the questionnaire, the probable impacts of diverse variables such as gender, GPA, daily use of internet/ computer and computer literacy over the student teachers' attitudes in general were analyzed. As the data were normally distributed and the variables were continuous, Pearson Correlation was run as to see the relation among variables. It is vital to mention here that correlation coefficient takes values between -1 and +1, which displays either negative or positive relation between variables. Also, once a correlation coefficient is higher than .70, this proves a strong relation. However, values lower than .20 mean a weak relation between variables. As for the coefficient value which is higher than .20 and is lower than .70, it is the sign of moderate relation between variables (Larson-Hall, 2010).

A Pearson Correlation test was computed in order to find out whether there was statistically significant relationship between the participants' attitudes towards computer assisted assessment and their demographic factors. As shown in Table

4.6, the findings of the correlation test revealed that there was no statistically significant relationship between any dimension of attitudes and gender. However, GPA had a weak negative relation with perceived ease of use ($r_{(97)} = -.213$, $p = .036$). These values showed that more successful participants might have difficulty in using the system with ease. The participants' daily usage of internet and computer displayed a weak positive relation with perceived ease of use ($r_{(97)} = .213$, $p = .037$). In this sense, it could be hold that the participants who spent much time on computer or on browsing the internet used the computer assisted assessment system easily. The last demographic factor, the participants' computer literacy levels, showed statistically significant relation with two dimensions of the questionnaire which were perceived ease of use and anxiety. According to the values belonging to the relation between computer literacy and perceived ease of use ($r_{(97)} = .244$, $p = .016$), it was seen that there was a moderate positive correlation. Therefore, having knowledge about how to use computers might ease the usage of computer assisted assessment systems. On the other hand, a moderate negative relation between computer literacy and anxiety was obtained ($r_{(97)} = -.343$, $p = .001$), which meant that anxiety level decreased in relation to the knowledge of computer use.

Table 4.6. Correlation Coefficients among Variables

		1	2	3	4	5	6	7	8	9
1. Gender	Pearson Correlation	1								
	Sig. (2-tailed)									
2. GPA	Pearson Correlation	.251 [*]	1							
	Sig. (2-tailed)	.013								
3. Internet Use	Pearson Correlation	.268 ^{**}	.040	1						
	Sig. (2-tailed)	.008	.694							
4. Computer Literacy	Pearson Correlation	.297 ^{**}	-.001	.289 ^{**}	1					
	Sig. (2-tailed)	.003	.996	.004						
5. Intention	Pearson Correlation	.155	-.100	.146	.138	1				
	Sig. (2-tailed)	.131	.331	.153	.177					
6. Perceives Usefulness	Pearson Correlation	.062	-.091	.137	.077	.893 ^{**}	1			
	Sig. (2-tailed)	.544	.374	.182	.452	.000				
7. Perceived Ease of Use	Pearson Correlation	.119	-.213 [*]	.213 [*]	.244 [*]	.782 ^{**}	.715 ^{**}	1		
	Sig. (2-tailed)	.244	.036	.037	.016	.000	.000			
8. Anxiety	Pearson Correlation	.150	.058	.154	-.343 ^{**}	.261 ^{**}	.174	.402 ^{**}	1	
	Sig. (2-tailed)	.142	.572	.133	.001	.010	.088	.000		
9. Computer Attitude	Pearson Correlation	.062	-.143	.155	.044	.439 ^{**}	.392 ^{**}	.563 ^{**}	.189	1
	Sig. (2-tailed)	.547	.163	.130	.669	.000	.000	.000	.064	

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

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

4.2.3. Results of Data Analysis for Research Question 3

Do the participants differ in their preferences of the exam mode?

In the present study, the experimental group was administered the computer-assisted testing whilst the control group took the traditional way of exam. This treatment might have impact on their preferences of exam mode for future courses. In order to define any difference between groups, Chi-square test was used as both variables were categorical (Larson-Hall, 2010, p. 135). Though the

data were normally distributed, Chi-square test had to be calculated because there is no parametric alternative to it.

Chi-square test value, $X^2(2, N=97) = 12.39$, proved the difference of the students' preferences up to their group, the experimental versus control groups. In addition, significant value was smaller than .05 ($p = .02 < .05$), which again asserted that there was a statistically significant difference between variables.

Table 4.7. Chi-Square Test Results

	<i>Value</i>	<i>df</i>	<i>Asymp. Sig. (2-sided)</i>
Pearson Chi-Square	12.399a	2	.002
Likelihood Ratio	12.815	2	.002
Linear-by-Linear Association	5.559	1	.018
N of Valid Cases	97		

p<0.05

In order to see what the students' choices in terms of exam mode for future courses were, Crosstabulation (Table 4.8) needed to be analyzed.

Table 4.8. Crosstabulation of the Students' Exam Preferences

		Exam Preference			Total	
		CAA	PBT	Both of them		
Exam Mode	PBT	Count	2	37	8	47
		% within Examtype	4%	79%	17%	100%
	CAA	Count	4	22	24	50
		% within Examtype	8%	44%	48%	100%
Total		Count	6	59	32	97
		% within Examtype	6%	61%	33%	100%

In total, 47 student teachers took the traditional exam. It seemed that they did not tend to change their exam mode since 37 of them stated that they preferred paper-based exams again. As for the experimental group's choices, it was quite concrete that the values were so close to each other. Of all participants in the experimental

group, 48% preferred to be offered both types of exam, but 44% preferred paper-based exam, and there were not so many participants who were in favor of using computer assisted assessment system.

4.2.4. Results of Data Analysis for Research Question 4

Is there any statistically significant difference in mean scores between the students who took a computer assisted test (the experimental group) and those who took a paper based achievement test (the control group)?

Mann Whitney U test was conducted in order to find out whether there was any statistically significant difference in mean scores between the students who took the computer assisted test and those who took traditional exam. As the dependent variable was continuous and the independent variable was categorical (the subjects were not the same in both groups) and the data related to the exam were not normally distributed, Mann Whitney U Test, non-parametric alternative to t-test, was used (Larson-Hall, 2010, p. 138).

The results (see Table 4.9) revealed that there was no significant difference in mean scores of the participants who took the computer assisted test and those who took paper based test, Z_U (954.500), $p = .111 > .05$. More clearly, once students take exams either on computer or on paper, their performance will not be affected in a good or bad way according to test administration mode.

Table 4.9. Mann Whitney U Test Results

<i>Group</i>	<i>N</i>	<i>Mean Rank</i>	<i>Sum of Ranks</i>	<i>Z_U</i>	<i>p*</i>
CAA	50	53.41	2670.50	954.500	.111
PBT	47	44.31	2082.50		

* $p > 0.05$

4.2.5. Results of Data Analysis for Research Question 5

Does the test administration mode influence the reliability and validity of computer-assisted versus paper-based tests?

While preparing and administering an exam, there are two backbones to which all teachers have to pay attention: reliability and validity. In this regard, both versions of the exam, computer-assisted and paper-based, are supposed to be reliable and valid in order to use them interchangeably and to benefit from computers as testing tools.

Firstly, split-half reliability method was calculated for each exam mode, and the reliability coefficients were given in Table 4.10. As seen, both computer-assisted and paper-based exams had significant level of reliability coefficients ($r_p = .75$, $r_c = .90$). Therefore, it was seen that computers had no drawbacks to be used as an alternative to traditional way of assessment.

Table 4.10. Reliability Coefficients of Computer-assisted and Paper-based Tests

<i>Test Mode</i>	<i>r</i>
<i>Paper-Based Test</i>	.756
<i>Computer-Assisted Test</i>	.903

Secondly, Spearman's rank order correlation, the non-parametric equivalence of Pearson-product moment correlation (Larson-Hall, 2010, p. 132), was conducted in order to reveal the correlation between both exam modes. While calculating the correlation coefficient, each exam item was assumed a case, not each student in this study. The findings (Table 4.11) displayed that there was a strong positive correlation ($r_p = .75$) between two test administration modes. This meant that the responses given by the test-takers in the experimental group were correlated with the ones given by the control group. Hence, the concurrent validity of the course achievement test was supplied.

Table 4.11. Spearman's Rank-order Correlation Results

		<i>CAA</i>	<i>PBT</i>
<i>Spearman's rho</i>	<i>Correlation Coefficient</i>	1.000	.894**
	<i>CAA Sig. (2-tailed)</i>	.	.000
	<i>N</i>	60	60
	<i>Correlation Coefficient</i>	.894**	1.000
	<i>PBT Sig. (2-tailed)</i>	.000	.
	<i>N</i>	60	60

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

4.3. Qualitative Data Analysis

In the present study, 26 randomly chosen participants were also interviewed in order to further explore their attitudes towards the computer assisted assessment system and to support the findings of quantitative research. The researcher thought that it would be of great importance to take into consideration the participants' thoughts and suggestions for the future studies as well as the present thesis. Findings of the qualitative data analysis are presented below under the last research question of the present study.

4.3.1 Results of Data Analysis for Research Question 6

What are the general attitudes and opinions of the participants toward the computer-assisted and paper-based tests?

Open ended questions were posed to the participants, and their responses were categorized under each research question. The interviewees' names were kept confidential; therefore, they were represented with PT (prospective teacher) and a number. In this sense, anonymity could be achieved.

Have you ever taken any formal computer-assisted assessment before? If so, how was your reaction(s) towards it?

Among all interviewees, only 5 students took a formal exam on computer before. When they talked about their experiences, 4 students claimed that they had taken a language test on computer, and they had really enjoyed it, but one student did not have so easy and enjoyable experience with the computer assisted assessment. PT5 explained it with these words:

"I took my math exam on computer when I was in America. It was my first experience, and well I had really though time on it. It was really scaring for me. No paper to calculate, no paper to take notes. Anyway, I did it."

As not many students had previous experience on computer-assisted assessment, other questions were of importance in order to show the prospective teachers' reactions to the system.

What did you like and/ or dislike about the computer assisted assessment?

After transcribing the audio files and analyzing the raw data, it was noticed that there were so many common points that the participants had touched on. What they liked and disliked about the system was categorized and shown below:

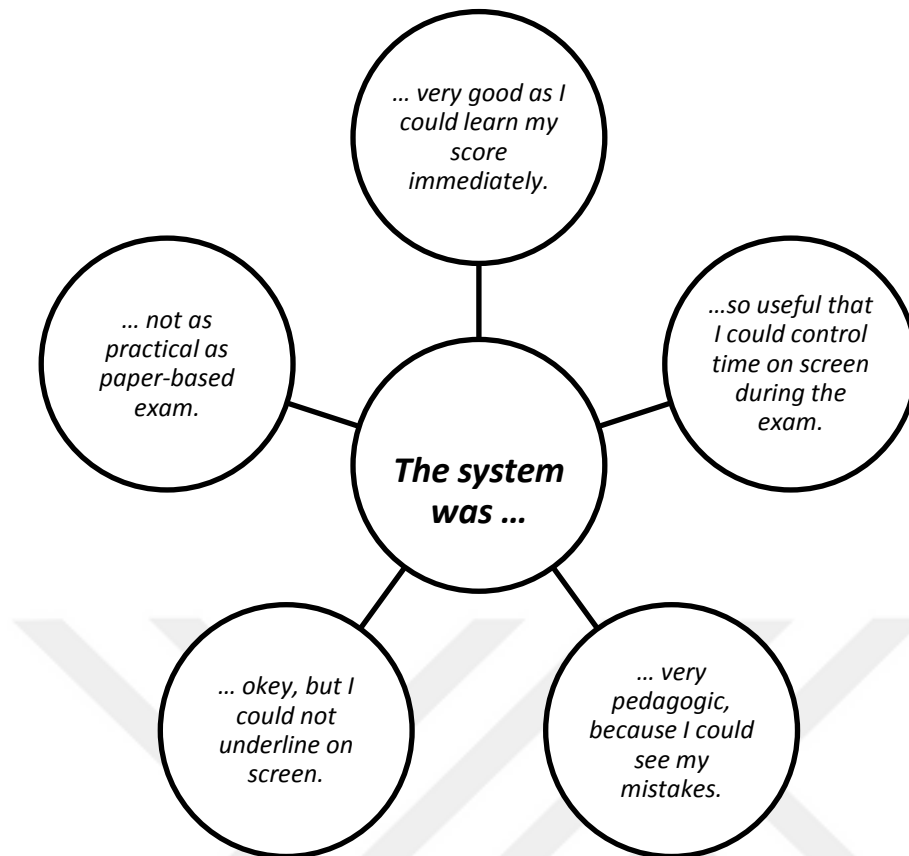


Figure 4.1. Comments of the Students

As seen, the students liked the system generally because of time control on screen and of pedagogical perspectives. PT16 uttered that despite the fear of using the system, its benefits outweighed.

"I was very stressed before I took the exam. I didn't think I would be able to do without paper and pencil. But the exam was very nice, it didn't turn out to be difficult at all. It is definitely more advantageous to take the test on the web. There is nothing to fear."

On the other hand, another participant was worried about being failure with this kind of testing systems and mentioned the opposite view:

"I want to see exam on paper, otherwise I cannot focus on the questions well enough. I need to underline some key words, but I cannot do it on computer. I may fail if I take all exams on computer." (PT8)

It was revealed that the participants had diverse thoughts on using computers as testing tools. As seen in the first interview item, many of the interviewees did not have any prior experience in terms of being tested on computers; hence, this was the first experience for them. It was quite natural to see different comments about the computer assisted system as personal differences were in due.

If you have to make a choice over a computer assisted exam and a traditional paper and pencil based exam, which one would you like to take? Why?

When the interviewees were asked to their choice, they showed tendency for paper-based exams; however, there were some participants who wanted to use computers as alternative testing tools. Despite the students' moderate attitude on accepting computer-assisted tests, they did not support the complete shift from the traditional testing to the technology-enhanced one. As for the reasons why they were against the changes over testing modes, many of them stated that they were using paper-based exams for many years, so they were accustomed to it. Some utterances of the students' who were in favor of continuing paper-based exams were given below:

"I prefer paper and pencil based exams, because it seems more concrete to me. I can touch paper and I can underline some words." (PT8)

"Of course I choose traditional exams, because we are accustomed to them." (PT10)

"I am not the kind of person who spends much time using the computer. So I prefer the traditional paper-based exam. It is necessary for me to underline the important places. I cannot do this on the computer." (PT7)

"... It was fine to take the exam online. I was able to control my time. I could answer the questions more quickly. But I prefer to take the traditional paper-and-pencil test, an old habit from the past. Holding the paper and underlining the important points increase my motivation." (PT4)

On the other hand, there were some students who both criticized and praised the system. What was common in the perceptions of these students was that they could use the computer assisted tests once exam was composed of multiple choice items, but writing on computer could be challenging. PT14 made the following comment:

"Computer assisted tests are good enough when exams include multiple-choice and/ or true-false items. However, I may not be efficient writing on computer. That's why, though I love the system for some aspects, I do not want to take all exams on computer. It can be offered as an alternative to us, so I can feel more confident."

Moreover, PT16 mentioned the preference by stating reasons. Accordingly, computer assisted assessment offered many advantages for the test-takers; therefore, they could be more motivated to the exam and could be more successful. This student stated that future courses should make use of computer technology for exams with these words:

"I like the system. First of all, I could control time on screen. Also, I saw my score just upon finishing the exam. I think that having a chance to revise my wrong answers was good. In this way, I can learn as well. Hence, exam is not only for assessing my knowledge, but also for teaching me. This is really good. And of course, administering exams on computer saves money. No need to pay for papers. Very economic way of assessing students. Despite the fear of using that kind of system, its benefits cannot be denied. I prefer taking exams on computers for future courses."

The prospective teachers showed differences in terms of their preferences of exam mode. Many of them were in thought of pursuing the traditional way of assessment while the rest participants supported the use of computer assisted assessment. The main reason of why they were against a radical shift was that they had been using paper based exam for years, so it might be hard for them to be tested on computers.

What advantages do you think a computer assisted exam has over a paper-based test?

The students were asked what they had experienced in the exam could be the good sides of the system; accordingly, all of them said immediate feedback as the first and profound advantage. Similarly, the students could see their mistakes, and this was uttered as the second most important side of computer-assisted exams. In this regard, the following comments were made:

"Actually, I did not expect to learn my score just upon the end of the exam, but it was fine. When exams are done on paper, it takes days or sometimes weeks to learn the score. This time not only did I learn it, but also I could see my mistakes. Therefore, now I know what topics I should study more. I wish other courses made use of it." (PT14)

"I could concentrate a lot while answering the questions on computer, so this helped me a lot and increased my performance during the exam. Also, I finished my exam, clicked the "finish" button and saw my score on the screen. This was excellent. I love the system a lot; undoubtedly, it has many benefits." (PT1)

What was common in these comments was that the students were satisfied with learning their grades and mistakes immediately.

What disadvantages do you think a computer assisted exam has over a paper-based test?

Following the question about the computer-assisted exam's advantages, the interviewees were posed a question in order to define what the disadvantages of the system were. More than half of the participants stated that reading on screen for a long time could be tiring. However, it is really surprising that normally they

spend so much time on computer, mostly for browsing the internet, and they do not feel tired. The following comment will clarify it:

"In general, I use my computer and mobile phone a lot, especially for social media. I spend much time on it, but I never feel tired. I guess it is because I enjoy and do not force my mind for anything. On the other hand, when I started my exam on computer, initially there was no problem; everything went well. However, after 15 or 20 questions, I started to feel tired since looking at screen for a long time and thinking the answer at the same time were a little bit hard, and my eyes got bleary." (PT6)

Another participant touched upon a different perspective that paper-based exams had advantage over computer assisted tests. Traditional exams make taking notes on paper while answering questions possible, and some of the students need to do it on exams. In this respect, students did not have any chance to take some notes that might help them remember important points of a topic and so on once technology integrated exam was administered. PT10 complained about the lack of paper and pencils in the exam:

"I am a visual learner, and I always take some short notes or draw charts or tables while studying, so I may need paper and pencil in the exam so as to write something which enables me to remember. I can do it in traditional exams, but cannot do it on computer. I like the system apart from it. Maybe a blank sheet and a pencil can be supplied to us. Also, open-ended questions, I mean writing long texts on computer, might be challenging for some students, because not all students are capable enough to write quickly and in a right way on computer."

PT1 criticized scoring factor on that sorts of exams. As the machine was responsible for grading the answers, no human factor was included.

"It was not the first time I took an exam on the web. I wasn't anxious about it, because I am accustomed to such types of tests. I was able to answer the questions more quickly, I didn't spend much time writing the answers on the answer sheet... The only thing that I can criticize on the computer assisted exam is that there is no human factor in scoring the test."

Actually, computer systems disappear subjective scoring; thus, this is an advantage, but it should be mentioned that before administering the exam, teachers have to define all possible answers to the system. Otherwise, even though a student gives a correct answer, but different from the alternative answers found on the system, computer will not accept it, which may turn the disadvantage side of computer assisted assessment.

Does using a computer-assisted system as an assessment tool save time?

Although the participants were posed questions about the computer-assisted exams' advantages and disadvantages, their performance in relation to time they had spent during the exam was the focal point of the present study. Nearly all of

them emphasized the time-saving aspect of computer-assisted tests. The following participants mentioned it with these words:

"The system did not require a lot of computer literacy, so it was quite easy to use. I read the question, decided the answer and clicked it. As for the fill-in-the blank part, it was easy to write on computer, too. I spent not much time to write answers on the answer sheet. Just one click was enough. Definitely, computer-assisted systems save our time." (PT1)

"...In addition to the benefits to us, I think the system is useful for teachers, as well. For example, I could learn my score just after the exam, so teachers do not need to grade all the papers." (PT9)

It is quite explicit that the participants are in favor of using computer assisted assessment tools because of their time-saving aspects. Not only teachers but also students can have great advantage over traditional assessment.

Do you think the computer-assisted assessment (specifically web tools) should replace the traditional paper-and-pencil assessment? Why or why not?

As the last part of the interview session, the participants' opinions about the change from paper-based assessment to computer-assisted assessment were asked. Though they thought that there were some good and bad sides of the system, there was a common point of view among all students: not only are technological tools available in order to teach students inside the classroom or to foster teaching process outside the classroom, but also they are available in order to be used as testing tools since technology is everywhere. The comment belonging to PT3 summarizes the topic:

"...well, to be honest, I was nervous when I learned that I would take the exam on computer, but it was not like that at all. Now I have experience upon taking computer-assisted exam, and I want to take it in the future as well, because we cannot survive without keeping updates in technology. And technology in education does not only cover teaching issues; assessment is a part of this world."

PT17, additionally, shed light upon the topic with these words:

"In time paper-based exams may be replaced by computer-assisted exams. Because I learned my score immediately after I took the exam on the web. I saw and learned my incorrect answers and those questions I didn't answer. This is very important for me. We used to wait for a very long time to learn our test scores. We couldn't learn anything about our incorrect answers and those questions we skipped to answer. We would likely make the same mistakes again. It is more different now, pedagogical."

As seen in the comments above, students are aware of the need to the change from traditional exams to technology integrated ones. Both computerized and paper based exams may be offered at the same time so as to make students

acquainted with computer-assisted assessment systems. After a certain period of time passes, a radical reform on test modes can be achieved, and teachers can make use of computers as assessment tools for all courses.



5. DISCUSSION OF THE FINDINGS

5.1. Introduction

With the widespread use of both computers and the internet, the state of education has changed to a great extent recently. Today, it is a must to benefit from the technological facilities not only to make students motivated towards learning, but also to enhance the assessment of the teaching and learning process. Computer-assisted assessment, which offers great advantages in terms of diverse perspectives over traditional exams, is a relatively recent test delivery mode, so there is not a great deal of scholarly literature related to its role in general. In this regard, the present study aimed to search the prospective English language teachers' attitudes and exam achievement toward computer assisted assessment systems. This part deals with the results of the research questions and the discussion by referring to the related studies done before.

5.2. Discussion of the Results

5.2.1. What are the prospective English teachers' attitudes towards computer-assisted assessment for their content knowledge course in general?

There are three belief components which are intention, perceived usefulness and perceived ease of use in the original TAM offered by Davies (1989). However, Alkış (2010) added two individual components, anxiety and computer attitude, to the model, and the present study made use of the latest version.

Upon giving ratings and making comments for each factor of the questionnaire separately, it is vital to combine them and to make connections to the previous related studies. In this study, the experimental group showed intention to computer assisted tests and they thought that the system had increased their performance or they would do better thanks to the system. Also, the experimental group assumed that they could learn how to navigate the system without any difficulty once the computer assisted tests were offered to them. Furthermore, when the related items under perceived ease of use were analyzed, it was seen that though the experimental group needed no support or assistance so as to handle the system, the control group might seek an assistant.

Alkış (2010) emphasizes perceived usefulness as the most motivating determinant, and perceived ease of use has positive relation to perceived usefulness and intention, which is similar to this study's findings. Accordingly, if students took exam without challenges, they would intent to use it again and believe in the efficiency of it. Also, Oz (2014) reveals similar findings with the present study; the student teachers had a positive attitude toward perceived ease of use, but were not sure with respect to their intent to take this sort of tests in the future, which was supported by the low values given to perceived usefulness by the student teachers.

The values show that students' previous experience on computer assisted testing system has no relation to their anxiety levels, which is consistent with Chin (1990), Dermo (2009) and Madsen (1991), yet they may be anxious toward the system since they are afraid of making a mistake not to be corrected. The anxiety factor may have tie to the intention as to take computer-assisted exams or usefulness of the system; thus, there may be a relation among their anxiety level, intention and perceived usefulness. Similarly, Alkış (2010) deduces once students have computer anxiety, their perception of ease, intention and perceived usefulness are affected negatively.

Overall, findings of both quantitative and qualitative data revealed that the participants favored the computer assisted assessment systems as they could learn their scores immediately. Also, they could get immediate feedback for their wrong and missing answers. Therefore, the pedagogical aspect of assessment, washback effect, was achieved, and the students were satisfied with this advantage of CAA. The values under computer attitude and perceived usefulness were supported by the interviewees' comments. Nevertheless, almost all of the students took their first experiences in the present study, yet the control group had no certain idea or experience on what CAA was. Thus, the missing of prior exposure to that kind of testing tools affected their preferences of exam mode administration for future courses; they did not intent to be tested with computers for all courses in the future, but they preferred the integration of both exam administration modes. This was similar with the values given to the related items of intention. Similarly, because of the lack of prior experience on CAA, the

students might show anxiety toward the system, which was again supported by the values under anxiety dimension of the questionnaire.

5.2.2. What is the relation between prospective English teachers' attitudes of computer assisted assessment and their demographic factors such as gender, GPA, daily use of computer and computer literacy?

The probable impacts of the students' gender, GPA, daily use of internet/ computer and computer literacy on their attitudes were also the focal point of the study. In this respect, the Pearson Correlation Coefficients among variables were measured, and it was found that gender had no statistically significant relation to the sub-scales of the questionnaire. Put differently, being a male or a female participant was not a prominent factor that affected the attitude toward the computer assisted system. This finding is in line with those of Chin (1990) and Dermo (2009) that report statistically no significant relation between students' gender and their ratings on the questionnaire, either.

However, there is a moderate positive relation between computer literacy and perceived ease of use; therefore, it can be stated that more proficient computer users navigate the system without any difficulty. Computer literacy, on the other hand, has a moderate negative correlation with anxiety, which asserts that proficient computer users show lower anxiety comparing to the students who are non-proficient. Also, grade point average and perceived ease of use have a moderate negative relation. Clearly, more successful participants may have difficulty in using the system with ease. And lastly, there is a moderate positive relation between perceived ease of use and daily use of internet and/ or computer, so it can be deduced that the participants who spend much time on computer can use the computer assisted testing system with great ease. Kirsch *et al.* (1998) and Taylor *et al.* (1998) wonder whether test-takers' prior computer knowledge have effect on their performances. The findings display that there is no significant effect of lack of computer knowledge over test-takers' scores (as cited in Choi, Kim and Boo, 2003).

5.2.3. Do the participants differ in their preferences of the exam mode?

As stated before, the experimental group was administered the computer-assisted assessment whilst the control group took the traditional way of exam in this study;

therefore, it was assumed that this treatment might cause differences on the student teachers' exam mode preferences for future courses. The related statistical data asserted the relation between the students' groups and their choices of exam modes. As for what the students' choices were, it was seen that the control group tended to pursue the classical way of exams, but there was a contradiction among the experimental group because nearly half of them preferred to be offered both types of exam, and the rest students wanted no change or no alternative to paper-and-pencil based tests. Moreover, there were not so many people who were in favor of using computer assisted assessment system, but it can still be offered as an alternative to students as some of them may use it.

The student teachers in the control group were strongly fond of paper-and-pencil based exams. The reason for this choice might be no prior experience on computer assisted assessment systems, so they had no certain idea about them. Undoubtedly, this situation had great impact over their exam mode choices for future courses and led the students to desire no change in the system. As to the student teachers in the experimental group, they were divided into two categories; one group with great interest in traditional exams and one group with moderate attitude to having computers as alternative assessment tools. At that point, the interviewees' comments should be taken into consideration as well. Some of them claimed that the system offered many advantages to them during the exam like time-control on screen, immediate feedback and score. In addition, the system helped them end their exam within a shorter period of time comparing to paper-based exams. However, majority of them had their first experiences in this study, so they were anxious about any radical shift from traditional to computer-assisted exam mode. This might be the reason of division among participants.

There are only three studies found in the related literature. In the study of Dermo (2009), it was seen that the students had a desire to take online test, because they were in thought of not only were they measured via online test, but also they could learn with this smart tool. This finding is consistent with the present study to a certain extent since some interviewees have made comments by touching washback effect of computer assisted assessment exams while giving their exam mode preferences in the future.

On the other hand, the findings are totally parallel with Oz (2014), which shows that the student teachers want to take neither fully computer assisted nor fully paper-based exams, but a combination of these two exam modes will trigger their success and motivation for future classes. Pawasauskas, Matson and Youssef (2014) also find that a majority of the participants want to take courses that are integrated with computer-assisted assessment systems, and this study is consistent with the present study, too.

5.2.4. Is there any statistically significant difference in mean scores between the students who took a computer assisted test (the experimental group) and those who took a paper based achievement test (the control group)?

Mann Whitney U test was calculated in order to see whether there was any statistically significant difference in mean scores between the experimental group and the control group. It was found that there was not any significant difference between the scores of the experimental group and the scores of the control group.

There are many studies done abroad and in Turkey about the equivalent exam scores which have been gotten on both exam modes, and the present study's finding is supported by many studies like Akdemir and Oğuz (2008), APA (1986), Bugbee (1996), Higgings, Russel and Hoffmann (2005), Logan (2015), Parshall and Kromrey (1993), Retnawati (2015), Russel, Golberg and O'Connor (2003), Stevenson and Gross (1991) and Yaman and Çağıltay (2010). On the other hand, Choi, Kim and Boo (2003) reached a little bit different result from these studies, because the students' listening scores were higher on computer-assisted test, but reading scores were higher on paper-and-pencil based exam. In this vein, it is vital to touch upon some related studies. For instance, Laborda (2010) states in his study that the visual cues presenting on computer create an authentic exam atmosphere for test-takers, and each examinee has an opportunity to listen to a text without being exposed to any external factors that may disturb them. Therefore, students' listening scores can go up. Within a similar perspective, Creed *et al.* (1987), Dillon (1994) and Muter *et al.* (1982) state that reading on screen is slower than on paper, and the accuracy of reading increases when students take their tests on paper.

On the contrary, it is really surprising and different from other studies that the students who took computer-assisted tests became more successful in Chin's study (1990), and Yağcı, Ekiz and Gelbal (2011) mentioned that the participants who took computer-assisted tests in their study succeeded 35% higher comparing to the participants who did not. As the related data analysis proves that there is no exam mode effect, computers can be adopted as alternative testing tools by teachers, because doubts about the affinity of students' scores both on paper and on computer have been eliminated.

5.2.5. Does the test administration mode influence the reliability and validity of the computer-assisted versus paper-based tests?

Validity and reliability are two crucial factors, so all teachers have to take into consideration both of them while preparing and administering an exam. In this sense, computer-assisted and paper-based exams are supposed to be reliable and valid in order to use them interchangeably and to benefit from computers as testing tools.

In order to see the reliability coefficients, split-half reliability method was run in this study. As both exam administration modes were done at the same time in the present study, they were supposed to supply the necessary concurrent validity values. The findings displayed that both computer-assisted and paper-based exams overlapped with respect to reliability values. This proved that there was no obstacle to use computer assisted assessment systems as an alternative to traditional way of assessment. In addition, Spearman's rank order correlation displayed that there was no significant difference between both exam types. It means that the responses given by the test-takers in the experimental group have strong tie with the ones given by the control group.

The results, in this sense, are consistent with Dermo (2009) and Siozos *et al.* (2009), but Chua (2012a, 2012b) concludes a different point from his studies since computer-assisted tests seem more reliable and offer more efficient testing environment for test-takers. This may be because of the fact that computers offer visual cues to test-takers, more authentic exam atmosphere can be achieved via computers, and test-takers have a change of listening to audio files individually with their earphones. Therefore, computer assisted tests serve in a desired way and test-takers can perform better on computers when they take some kinds of

tests such as language skill-based tests (listening items or reading items with some graphs or pictures).

Although Chua's studies and the present study's findings are different from each other, it should be mentioned that computer assisted testing or web based testing can be benefitted by educators without any doubts since they have similar values in terms validity and reliability. Despite higher reliability values in Chua's study, this is a good indicator of computers as testing tools. Likewise, Delen (2015) aims to put forward how to increase the validity and reliability of computer-assisted tests. In this sense, optimum item response time for each question was shown on the screen when the participants were busy with answering the exam items. The findings reveal that if students are offered optimum item response time, more valid and reliable tests will be achieved rather than paper-based tests. Delen's study does not have totally same goals with the present study; however, paying attention to validity and reliability in a test is so crucial.

5.3. Conclusion

In general, in Turkey, it seems hard to make a radical shift from traditional testing to technology-based one, but students are capable enough to handling and use computer-assisted assessment systems. Therefore, initially computers can be offered as alternative testing tools in addition to traditional tests, and day by day students are getting acquainted with the smart testing system. Then, universities and schools can start to implement totally computer-assisted tests. Also, CAA offers reliable and valid tests, which are seen two backbones in assessment, and supplies many advantages to teachers, institutions and students.

6. CONCLUSION

6.1. Introduction

This chapter presents an overview of the study and some pedagogical implications as well as limitations. It also makes some suggestions for future studies.

6.2. Overview of the Study

Assessment is an indispensable part of education, and teachers are supposed to pay particular attention to some criteria such as validity (whether the test actually measures what it is supposed to measure), reliability (which refers to the consistency level of a test or a measurement instrument done at a different time and place) and practicality (Hughes, 2003, pp. 33-34). In addition, assessment has a great impact on teaching and learning process, which is known as “washback effect” in the literature. Certainly, it is really crucial to prepare tests, which are valid, reliable and practical, to provide a relaxed exam environment for test-takers and to grade them in an objective way. Therefore, assessment may be viewed as a challenging task and even as a burden generally for teachers and teacher educators.

In this sense, with the introduction of computers into education, technology provides a number of opportunities in the assessment of learning and teaching outcomes in education. One recent popular trend is computer-assisted assessment that offers many advantages like speed, flexibility, efficiency, fall in the challenges involved in test security, creation randomized questions from large item pools and immediate feedback.

With the growing importance of technology in education, the present study aimed to investigate the role of computer-assisted assessment on the perceptions and course achievement performance of prospective English language teachers, because not only were teachers’ thoughts of importance in order to define the pros and cons of computer-assisted testing system, but also students’ views needed to be taken into consideration. Furthermore, gender differences, grade point average, prior knowledge of computer literacy, daily usage of computers were among the challenging questions that needed to be answered. Any difference among the students’ exam mode preferences for future courses, their mean scores in relation

to test mode effect and the validity and reliability coefficients of both exam modes were tried to be searched. In the pursuit of these aims, 97 student teachers who were in the department of English Language Teaching at Hacettepe University were randomly assigned to experimental and control groups. In order to collect data, two types of instruments were used: computer-assisted assessment questionnaire and interview.

Computer-assisted assessment questionnaire has three belief components which are intention, perceived usefulness and perceived ease of use and two individual components which are anxiety and computer attitude. According to the results of the study, the participants in the experimental group seemed to have intent to computer-assisted tests. On the other hand, the student teachers in the control group showed no interest or intention toward the system. Likewise, the majority of the students in the control group did not favor the system. However, the experimental group believed in the usefulness of the system. Put differently, the system might increase their performance or they would do better thanks to it. Hence, it can be said that intention and perceived usefulness have a relation. In addition, many students in the experimental group assumed that they could learn how to navigate the system without any difficulty if computer-assisted tests were offered to them, which referred to perceived ease of use dimension. Alkış (2010) emphasized perceived usefulness as the most motivating determinant, and perceived ease of use had positive relation to perceived usefulness and intention, which was similar to this study's findings. Likewise, Oz (2014) revealed similar findings with the present study. The findings displayed that students' previous experience on computer assisted testing system had no relation to their anxiety levels, and so did Chin (1990), Dermo (2009) and Madsen (1991).

The students' attitudes toward computer assisted tests may change up to their gender, GPA, daily use of internet/ computer and computer literacy. The findings of the present study displayed that gender had statistically no significant relation to the sub-scales of the questionnaire. Chin (1990) and Dermo (2009) found no statistically significant relation between students' gender and their ratings on the questionnaire, either. However, there was a moderate positive relation between computer literacy and perceived ease of use, which meant that more proficient computer users used the system without challenges. Computer literacy, on the

other hand, had a moderate negative correlation with anxiety, so proficient computer users showed lower anxiety comparing to the students who were non-proficient. In addition, grade point average and perceived ease of use had a moderate negative relation, and lastly, there was a moderate positive relation between perceived ease of use and daily use of computer. Thus, it can be concluded that students who spend much time on computer can use computer assisted assessment system with great ease. Within a similar perspective, Kirsch *et al.* (1998) and Taylor *et al.* (1998) found no significant effect of lack of computer knowledge over test-takers' scores (as cited in Choi, Kim and Boo, 2003).

When it comes to compare the groups' test administration mode choices, the student teachers in the control group were very fond of taking paper-and-pencil based tests. A distinction among the participants in the experimental group was seen as nearly half of them preferred to be offered both types of exam and the rest students desired no change. Dermo (2009) found the students' desire to take online tests. On the other hand, the findings of this study totally overlapped with Oz (2014) which also showed that the student teachers wanted to take neither fully computer assisted nor fully paper-based exams, but a combination of these two exam modes would trigger their success and motivation for future classes.

Whether students' scores showed any changes according to test administration mode, either on computer or on paper, or not was also shed light in this study, and no significant difference was obtained from the related findings. Thus, equivalent exam scores can be achieved without taking into consideration the exam mode effect (Akdemir and Oğuz, 2008; APA, 1986; Bugbee, 1996; Higgings, Russel and Hoffmann, 2005; Logan, 2015; Parshall and Kromrey, 1993; Retnawati, 2015; Russel, Golberg and O'Connor, 2003; Stevenson and Gross, 1991; Yaman and Çağıltay, 2010). On the contrary, Choi, Kim and Boo (2003) reached a little bit different result: listening scores were higher on computer-assisted exam version, but reading scores were higher on paper-and-pencil based exam. Within a similar perspective, Creed *et al.* (1987), Dillon (1994) and Muter *et al.* (1982) claimed that reading on screen was slower than on paper, and Laborda (2010) emphasized the computers' benefits for listening items.

Validity and reliability are two crucial factors that all teachers have to pay attention. In this regard, computer-assisted and paper-based exams are supposed to be

reliable and valid in order to use them interchangeably and to benefit from computers as testing tools. It was found that both exam versions overlapped with respect to reliability and validity values, which showed that computers can be used as alternative assessment tools to traditional exams. The results are consistent with Dermo (2009) and Siozos *et al.* (2009), but Chua (2012a, 2012b) concludes a different point from his studies since computer-assisted tests seem more reliable and offer more efficient testing environment for test-takers.

All in all, although it seems hard to make a radical shift from traditional testing to technology-based one, especially within the context of Turkey when the participants' attitude ratings are taken into consideration, students are capable enough to handling computer-assisted test. Therefore, it is quite logical to utilize computers as testing tool alternatives and to make students acquainted with this smart system. Then, universities and schools can start to implement totally computer-assisted exams. In addition, it is crucial to mention that computer-assisted exams have no obstacle or drawbacks in term of offering reliable and valid tests, which are seen two backbones of tests, many advantages supplied by system to teachers, institutions and students and making students get equal and comparable scores to paper-based tests.

6.3. Pedagogical Implications

This study aimed to find out the student teachers' attitudes toward computer-assisted assessment system. However, not only their perceptions to the system, but also efficiency of this type of tests were aimed to be searched. The findings revealed many pedagogic implications, and now each of them is explained.

No matter which test design is used, testing has to be based on some principles such as validity, reliability, practicality and washback effect (or backwash effect). Once a test has these principles, it proves the efficiency of it. This study compared computer-assisted and paper-based modes of the same test. No significant difference between them was recognized. Moreover, computers offer very rich test content, especially for language tests, and supplies the base for communicative language testing (Brown, 2004; Choi, Kim and Boo, 2003; Noyes and Garland, 2008). Clearly, visual cues that are shown on screen during listening tests (Laborda, 2010) or reading tests make the exam atmosphere much more

authentic, which is the key component of communicative language testing. Therefore, computer-assisted exams are said to be valid as they serve the test's aim in a desired way.

CAA also gives immediate feedback for incorrect and missing answers (Alderson, 2000; Cohen, 2001; Noijons, 2013; Yunxiang *et al.*, 2010), so students have a chance to learn their deficient points and to focus on these areas. Contrary to traditional testing, which takes a long time to announce results and seems a burden for teachers, and most of the time, it is impossible for teachers to give enough feedback for each learner about their mistakes on exam items; computer-assisted testing makes delivering test scores and giving feedback just upon completing the test possible (Alderson, 2000, p. 595), because giving feedback right after any mission done has a crucial and meaningful impact on learning, which is useful for pedagogical purposes (Roever, 2001, p. 85), and assessment is done for both grading students and measuring teaching process, which refers to washback effect.

Machine-based or computer-based scoring removes the burden over teachers, and subjectivity on scoring disappears. In addition, special programs are available in order to aid test design, item editing, piloting and having an item pool, which again serves for the principles of effective tests.

Another implication is about language tests such as TOEFL, IELTS, PTE and Turkish national foreign language proficiency exam (YDS) that address to instrumental purposes as the test-takers can use the results of these aforementioned tests in order to find jobs, to get promoted and so on. Today, large-scale exams like TOEFL and IELTS are done via WBT and YDS has started to be implemented via WBT recently. Hence, getting students to become familiar with CAA with classroom applications will be fruitful.

Lastly, it should be mentioned that teaching and assessment procedures are inseparable, and today individualized learning appears in relation to technology-enhanced learning tools. Also, Krashen (1982) mentions that once a learner is taught beyond his/her current level, teaching is done more effectively. The latest improvements in computer assisted assessment have given way to measure each examinee's level separately: computer adaptive testing, which is driven by Item

Response Theory (Drasgow, 2002, p. 4). Computer adaptive tests support Krashen's theory, which gives way to assess each learner's knowledge not with same questions for all, but with different questions depending on large amount of information (Sokolik, 2001).

6.4. Limitations of the Study

The research reached its aims, but there were certainly some limitations. First of all, only 97 student teachers in the department of English language teaching were included. More students in the same department from different universities or more students from a variety of departments could have participated in the study, so the results could be easily generalized to higher education system in Turkey. Secondly, the study was only focused on tertiary level students, but high school, secondary school and/ or even primary school students could be covered in these studies. Thirdly, the participants took the computer assisted test only one time, for their midterm exam, but using this system during one semester or during whole year may give more sensitive results. And lastly, computer adaptive testing is the paradigm of the 21st century; however, the researcher couldn't use it, but computer-assisted testing system was utilized.

6.5. Suggestions for Further Research

Now that there are some limitations in the study, further studies can focus on these points. First of all, the researcher used computer-assisted assessment system, but computer-adaptive tests are the contemporary exam modes in this century. Hence, future researchers can make use of these tests.

The participants in this study took the exam on computer only once for their midterm exam. Future studies can last during one year or one semester. In this sense, students' attitudes toward computer-assisted system at the beginning of year and at the end of year can be compared, and the effect of continuous exposure to the system can be analyzed. In addition, students from different departments or from different universities can be included, because as the number of samples go up, the findings can be generalized in a salient way and the small size effect can be achieved (Albers, Boeve and Meijer, 2015).

As a last suggestion, scoring effect can be given. Immediate feedback on computer systems is given at two different times: either after all items are

answered or after each incorrect or missing answer is given. And if teachers allow, students can answer the same question for the second time when it is missing or incorrect. Here arises a question. Which answer should be graded? Callı *et al.* (2003) think that the second response should be scored, because learning has occurred after getting feedback. Still, this point needs to be shed light on through further studies.

6.6. Conclusion

It is vital to benefit from technology at all levels of education and for a variety of goals, because without it, nothing is achieved completely. Students today spend most of their time on computers and on mobile phones, and they can learn a lot on their own with these smart tools. Likewise, teachers can be global thanks to technological inventions. This study focused on only computer assisted assessment and its pros and cons as assessment tools. It was seen that computers offered many advantages as testing aids, so educators should turn their way to use them, by bearing in mind the limitation of this study and by focusing on the suggestions given, not only to teach but also to assess students.

REFERENCES

- Akdemir, O., & Oğuz, A. (2008). Computer-based testing: An alternative for the assessment of Turkish undergraduate students. *Computers & Education*, 51, 1198-1204.
- Akkoç, H. (2012). Bilgisayar destekli ölçme-değerlendirme araçlarının Matematik öğretimine entegrasyonuna yönelik hizmet öncesi eğitim uygulamaları ve Matematik öğretmen adaylarının gelişimi. *Turkish Journal of Computer and Mathematics Education*, 3 (2), 99-114.
- Al-Saleem, S.M., & Ullah, H. (2014). Security considerations and recommendations in computer-based testing. *The Scientific World Journal*, 2014, 1-7.
- Albers, C.J., Boeve, A.J., & Meijer, R.B.(2015). A critique to Akdemir and Oguz (2008): Methodological and statistical issues to consider when conducting educational experiments. *Computers & Education*, 87, 238-242.
- Alkış, N. (2010). *Identifying factors that affect students' acceptance of web-based assessment tools within the context of higher education*. Unpublished master's thesis. METU, Ankara.
- Alderson, J.C. (2000). Technology in testing: The present and the future. *System*, 28, 53-603.
- American Psychological Association (1986). *Guidelines for computer-based tests and interpretations*. Washington, DC: Author.
- Antoun, J., Nasr, R., & Zgheib, N. K. (2015). Use of technology in the readiness assurance process of team based learning: Paper, automated response system or computer based testing. *Computers in Human Behavior*, 46, 38-44.
- Bachman, L. F., & Palmer, A. S. (2000). *Language testing in practice*. Oxford: Oxford University Press.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması* [Measurement in education and psychology: Classical test theory and its application] Ankara: OSYM Yayınları.
- Blerkom, M. L. V. (2009). *Measurement and statistics for teachers*. New York, NY: Routledge.
- Bridgeman, B. (2009). Experiences from large-scale computer-based testing in the USA. In F. Scheurmann & J. Björnsson (Eds.). *The transition to computer based assessment: New approaches to skills assessment and implications for large-scale testing*, 39-44. Luxembourg: Office for Official Publications of the European Communities.
- Brown, H. D. (2000). *Principles of language learning and teaching*. New York, NY: Longman.
- Brown, H.D. (2004). *Language assessment: Principles and classroom practices*. White Plains, NY: Pearson Education.

- Brown, H. D. & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices*. White Plains, NY: Pearson Education.
- Brown, J. D. (1996). *Testing language programs*. Upper Saddle River, New Jersey: Prentice Hall Regents
- Brown, J. D. (2004). Research methods for applied linguistics: Scope, characteristics, and standards. In A. Davies & C. Elder (Eds.). *The handbook of applied linguistics*, 476-500. Oxford: Blackwell.
- Brusilovsky, P., & Miller, P. (1999). *Web-based testing for distance education*. Webnet 99 World conference on the WWW, Hawaii, USA, 24-30 October 1999.
- Bugbee, A. C. (1996). The equivalence of paper-and-pencil and computer-based testing. *Journal of Research on Computing in Education*, 28 (3), 282-299.
- Bull, J., & McKenna, C. (2004). *Blueprint for computer-assisted assessment*. London: RoutledgeFalmer.
- Büyüköztürk, S. (2011). *Sosyal bilimler için veri analizi el kitabı* (14th ed.) [Handbook of data analysis in social sciences]. Ankara: Pegem Akademi.
- Callı, İ., Torkul, O., ve Taşbaşı, N. (2003). İnternet destekli öğretimde kullanılmak üzere web erişimli veri tabanı yönetim sistemiyle ölçme ve değerlendirme sistemi tasarımı. *TOJET*, 2 (3), 108-115.
- Cambridge ESOL (2008). *Research notes*. [Online: <http://www.cambridgeenglish.org/images/23151-research-notes-32.pdf>, retrieved in August, 2015].
- Carton, E.K., Carton, A. S., & Dandonoli, P. (1991). Developing a computer-adaptive test of French reading proficiency. In P. Dunkel (Ed.) *Computer-assisted language learning and testing*, 259-284. New York, NY: Newbury House.
- Chapelle, C. A. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing and research*. Cambridge: Cambridge University Press.
- Chapelle, C. A., & Douglas, D. (2006). *Assessing language through computer technology*. Cambridge: Cambridge University Press.
- Cheng, K. (2009). Computer-based language testing: Perspectives of the past and the future. *4th International Conference on Computer Science and Education*, Nanning, China, 25-28 July 2009.
- Chin, C. H. L. (1990). *The effect of computer-based tests on the achievement, anxiety and attitudes of grade 10 science students*. Unpublished master's thesis. The University of British Columbia, Vancouver.
- Choi, I., Kim, K.S., & Boo, J. (2003). Comparability of a paper-based language test and a computer-based language test. *Language Testing*, 20 (3), 295-320.
- Chua, Y. P. (2012a). Effect of computer-based testing on test performance and testing motivation. *Computers in Human Behavior*, 28, 1580-1586.

- Chua, Y. P. (2012b). Replacing paper-based testing with computer-based testing in assessment: Are we doing wrong? *Procedia-social and Behavioral Sciences*, 64, 655-664.
- Cisar, S. M., Radosav, D., Markoski, B., Pinter, R., & Cisar, P. (2010). *New Possibilities for Assessment through the Use of Computer Based Testing*. 8th International Symposium on Intelligent Systems and Informatics, Serbia, 10-11 September 2010 .
- Cohen, A. D. (2001). Second language assessment. In M. Celce-Murcia (Ed.). *Teaching English as a second or foreign language*, 515-534 (3rd ed). Boston, MA: Heinle&Heinle.
- Comert, M. (2008). *Bireye uyarlanmış bilgisayar destekli ölçme ve değerlendirme yazılımı geliştirilmesi*. Unpublished master's thesis. Bahçeşehir Üniversitesi, Istanbul.
- Conejo, R., Guzman, E., Millan, E., Trella, M., Perez-De-La-Cruz, J. L., & Rios, A. (2004). SIETTE: A web-based tool for adaptive testing. *International Journal of Artificial Intelligence in Education*, 14, 1-33.
- Corbin, J. & Strauss, A. (1990). Grounded theory research: Procedures, canons and evaluative criteria. *Qualitative Sociology*, 13(1), 3-21.
- Creed, A., Dennis, I., & Newstead, S. (1987). Proof-reading on VDUs. *Behaviour and Information Technology*, 6(1), 3-13.
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research (4th ed.)*. Boston, MA: Pearson.
- Davies, A. (1983). Computer-assisted language testing. *CALICO Journal*, 1(5), 41-43.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35(8), 982-1003.
- Delen, E. (2015). Enhancing a computer-based testing environment with optimum item response time. *Eurasia Journal of Mathematics, Science and Technology Education*, 11(6), 1457-1472.
- Demircioğlu, G. (2007). Geçerlik ve güvenirlik [Validity and reliability]. In E. Karip, (Ed.) *Ölçme ve değerlendirme*, 51-79 [Assessment and evaluation]. Ankara: Pegem Yayıncılık.
- Dermo, J. (2009). E-assessment and the student learning experience: A survey of student perceptions of e-assessment. *British Journal of Educational Technology*, 40 (2), 203-214.
- Dillon, A. (1994). *Designing usable electronic text: Ergonomic aspects of human information usage*. London: Taylor & Francis.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.

- Drasgow, F. (2002). The work ahead: A psychometric infrastructure for computerized adaptive tests. In C. N. Mills, Potenza, M. T., Fremer, J. J., Ward, W. C. (Eds.) *Computer-based Testing*, 1-40. Mahwah, NJ: Lawrence Erlbaum Associates Pub.
- Dunkel, P.A. (1999). Considerations in developing or using second/foreign language proficiency computer-adaptive tests. *Language Learning and Technology*, 2 (2), 77-93.
- Ekane, W. W., & Maiken, D. (1997). The English vocabulary-technology connection: A third world experience. *Educational Media International*, 34 (3), 125-128.
- Erdogdu, B. (2009). *Computer based testing evaluation of question classification for computer adaptive testing*. Unpublished master's thesis. Bahçeşehir Üniversitesi, Istanbul.
- Fisher, D., & Frey, N. (2007). *Checking for understanding*. Alexandria, Virginia: ACSD
- Flaughner, R. (2000). Item banks. In H. Wainer, N. J. Dorans, D. Eignor, R. Flaughner, B. F. Green, R. J. Mislevy, L. Steinberg, & D. Thissen (Eds.), *Computerized adaptive testing: A primer*, 37-59. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Folk, V. G., & Smith, R. L. (2002). Models for delivery of CBTS. . In C. N. Mills, Potenza, M. T., Fremer, J. J., Ward, W. C. (Eds.), *Computer-based testing: Building the foundation for future assessments*, 41-66. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Freeman, D. L., & Anderson, M. (2011). *Techniques and principles in language teaching*. Oxford: Oxford University Press.
- Frey, N., & Fisher, D. (2011). *The formative assessment action plan*. Alexandria, Virginia: ACSD
- Garcia, M. R. & Arias, F.V. (2000). A comparative study in motivation and learning through print-oriented and computer-oriented tests. *Computer-assisted Language Learning*, 13 (4), 457-465.
- Gardner, J. (2012). *Assessment and learning*. London: Sage Publications, Inc.
- Gronlund, N. E. (1998). *Assessment of student achievement*. Boston: Allyn and Bacon.
- Guzman, E., & Conejo, R. (2005). Self-assessment in a feasible, adaptive web-based testing system. *IEEE Transactions on Education*, 48 (4), 688-695.
- Harlen, W. (2012). On the relationship between assessment for formative and summative purposes. In J. Gardner, (Ed.). *Assessment and learning*, 87-102. London: Sage Publications, Inc.
- Hatch, E., & Lazaraton A. (1991). *The research manual: Design and statistics for applied linguistics*. New York, NY: Newbury House.
- Heaton, J. B. (1988). *Writing English language tests*. Harlow, Essex: Longman.
- Henning, G. (1991). Validating and item bank in a computer-assisted or computer-adaptive test. In P. Dunkel (Ed.), *Computer-assisted language learning and testing*, 209-222. New York, NY: Newbury House.

- Higgings, J., Russell, M., & Hoffmann, T. (2005). Examining the effect of computer-based passage presentation on reading test performance. *Journal of Technology, Learning and Assessment*, 3 (4), 3-35.
- Hudspeth, D. (1997). Testing learner outcomes in web-based instruction. In B. H. Khan (Ed.), *Web-based instructions*, 353-356. Englewood Cliffs: Educational Technology Publications.
- Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge, England: Cambridge University Press.
- Hymes, D. H. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.) *Sociolinguistics: selected readings*, 269-293. Harmondsworth: Penguin.
- Kan, A. (2013). Ölçme araçlarında bulunması gereken nitelikler [Qualifications to be found in assessment tools]. In H. Atılğan (Ed.), *Eğitimde ölçme ve değerlendirme*, 23-80 (6th ed.). [Assessment and evaluation in education]. Ankara: AnıYayıncılık.
- Kearsley, G. (1996). The World Wide Web: Global access to education. *Educational Technology Review*, 5, 26-30.
- Khan, B. H. (1997). Web-based instruction (WBI): What is it and why is it? In B. H. Khan (Ed.), *Web-based instructions*, 5-18. Englewood Cliffs: Educational Technology Publications.
- Kirsch, I., Jamieson, J., Taylor, C., & Eignor, D. (1998). *Computer familiarity among TOEFL examinees*. (TOEFL Research report no:59). Princeton, NJ: Educational Testing Service.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Kress, G. (2009). Assessment in the perspective of a social semiotic theory of multimodal teaching and learning. In C. W. Smith, & J. Cumming. *Educational assessment in the 21st century*, 19-41. NewYork: Springer B.V.
- Laborda, J. G. (2010). Contextual clues in semi-direct interviews for computer assisted language testing. *Procedia Social and Behavioral Sciences*, 2, 3591-3595.
- Laborda, J. G., & Litzler, M. F. (2011). Constraints in teacher training for computer assisted language testing implementation. *International Education Studies*, 4 (2), 13-17.
- Laborda, J. G., & Royo, T. M. (2009). Training senior teachers in compulsory computer-based language tests. *Procedia Social and Behavioral Sciences*, 1, 141-144.
- Laborda, J.G., Santiago, M.L., Juan, N.O., & Alvarez, A.A. (2014). Communicative language testing: Implications for computer based language testing in French for specific purposes. *Journal of Language Teaching and Research*, 5(5), 971-975.
- Larsen-Freeman, D., & Anderson, M. (2011). *Techniques and principles in language teaching* (3rd ed.). Oxford, England: Oxford University Press.
- Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. Abingdon, Oxon: Routledge.

- Lee, M. K. O., Cheung, C. M. K., & Chen, Z. (2005). Acceptance of internet-based learning medium: The role of extrinsic and intrinsic motivation. *Information and Management*, 42, 1095-1104.
- Lilley, M., Barker, R., & Britton, C. (2004). The development and evaluation of a software prototype for computer-adaptive testing. *Computers and Education*, 43, 109-123.
- Linden, W. J. (2002). On complexity in CBT. . In C. N. Mills, Potenza, M. T., Fremer, J. J., Ward, and W. C. (Eds.), *Computer-based testing: Building the foundation for future assessments*, 89-102. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Linden, W. J., & Glas, G. A. W. (2002). *Computer-adaptive testing: Theory and Practice*. NewYork: Kluwer Academic Publishers.
- Logan, T. (2015). The influence of test mode and visuospatial ability on mathematics assessment performance. *Mathematics Education Research Journal*, 27, 423-441.
- Luecht, R. M. (2001). *Challenges of web-based assessment*. Annual meeting of the national council on measurement in education, Seattle, April 2001.
- Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Madsen, H.S. (1983). *Techniques in testing*. New York, NY: Oxford.
- Madsen, H.S. (1986). Evaluating a computer-adaptive ESL placement test. *CALICO*, 4(2), 41-50.
- Madsen, H. S. (1991). Computer-adaptive testing of listening and reading comprehension. In P. Dunkel(Ed.) *Computer-assisted language learning and testing*, 237-257. New York, NY: Newbury House.
- McGough, J., Mortensen, J., Johnson, J., & Fadali, S. (2001). *A web based testing system with dynamic question generation*. 31st ASEE/ IEEE frontiers in education conference, Reno, 10-13 October 2001.
- McNamara, T. (2000). *Language testing*. Oxford: Oxford University Press.
- Muter, P., Latremouille, S. A., Treurniet, W. C., & Beam, P. (1982). Extended reading of continuous text on television screens. *Human Factors*, 24, 502-508.
- Nichols, G. W. (1997). Formative evaluation of web-based instruction. In B. H. Khan (Ed.), *Web-based instructions*, 369-374. Englewood Cliffs: Educational Technology Publications.
- Noijons, J. (1994). Testing computer assisted language testing: Towards a checklist for CALT. *CALICO Journal*, 12 (1), 37-58.
- Noyes, J. M., & Garland, K. J. (2008). Computer- vs. paper-based tasks: Are they equivalent? *Ergonomics*, 51(9), 1352-1375.
- Ong, C. S. & Lai, J. Y. (2006). Gender differences in perceptions and relationships among dominants of e-learning acceptance. *Computers in Human Behavior*, 22, 816-829.

- Oz, H. (2014). Pre-service English teachers' perceptions of web-based assessment in a pedagogical content knowledge course. *Procedia - Social and Behavioral Sciences*, 141, 45-58.
- Ozmen, K. S., ve Balçıkanlı, C. (2012). Yabancı dil eğitiminde web tabanlı ölçme ve değerlendirme. In A. Sarıçoban ve Z. M. Tavil (Eds.), *Yabancı dil öğretiminde öğretim teknolojileri ve material tasarımı* (pp. 229-251). Ankara: Anı Yayıncılık.
- Park, S. Y. (2009). An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning. *Educational Technology & Society*, 12 (3), 150-162.
- Parshall, C. G., & Kromrey, J. D. (1993). *Computer-based versus paper-and-pencil testing: An analysis of examinee characteristics associated with mode effect*. Annual meeting of the American educational research association, Atlanta, GA, April 1993.
- Parshall, C. G., Spray, J. A., Kalohn, J. C., & Davey, T. (2002). *Practical considerations in computer based testing*. Verlag, NewYork: Springer.
- Pawasauskas, J., Matson, K.L., & Youssef, R. (2014). Transitioning to computer based testing. *Currents in Pharmacy Teaching and Learning*, 10, 289-297.
- Ployhart, R. E., Weekley, J. A., Holtz, B. C., & Kemp, C. (2003). Web-based and paper-and-pencil testing of applicants in a proctored setting: Are personality, biodata and situational judgment tests comparable? *Personnel Psychology*, 56, 733-752.
- Ravid, R. (2011). *Practical statistics for educators* (4th ed.) Plymouth, UK: Rowman & Littlefiel.
- Reckase, M. D. (1989). Adaptive testing: The evaluation of a good idea. *Educational Measurement: Issues and Practice*, 8(3), 11-15.
- Retnawati, H. (2015). The comparision of accuracy scores on the paper and pencil testing versus computer-based testig. *TOJET*, 14(4), 135-142.
- Richards, J. C. (2006). *Communicative language teaching today*. Cambridge, England: Cambridge University Press.
- Richards, J. C., & Rodgers, T. S. (2001) *Approaches and methods in language teaching* (2nd ed.). Cambridge, England: Cambridge University Press.
- Robinson, G.L. (1991). Effective feedback strategies in CALL. In P. Dunkel, (Ed.) *Computer-assisted language learning and testing*, 155-167. New York, NY: Newbury House.
- Roever, C. (2001). Web-based language testing. *Language Learning and Technology*, 5 (5), 84-94.
- Rowe, J. A. (2007). Computer-based testing for the nursing discipline: A new frontier. *Journal of Emergency Nursing*, 33 (2), 164-168.
- Russell, M., Goldberg, A., & O'conner, K. (2003). Computer-based testing and validity: A look back into the future. *Assessment in Education: Principles, Policy & Practice*, 10 (3), 279-293.

- Saade, R., & Bahli, B. (2005). The impact of cognitive absorption on perceived usefulness and perceived ease of use in on-line learning: An Extension of the technology acceptance model. *Information & Management*, 42, 317-327.
- Salaberry, M.R. (2001). The use of technology for second language learning and teaching: A retrospective. *The Modern Language Journal*, 85(1), 39-56.
- Scheerens, J., Glas C., & Thomas, S. M. (2005). *Educational evaluation, assessment, and monitoring: A systemic approach*. Lisse: Swets & Zeitlinger B.V.
- Seidman, I. (2006). *Interviewing as qualitative research: A guide for researchers in education and the social sciences (3rd ed.)*. New York, NY: Teachers College.
- Semerçi, Ç., ve Bektaş, C. (2005). İnternet temelli ölçmelerin geçerliliğini sağlamada yeni yaklaşımlar. *TOJET*, 4 (1), 130-134.
- Siozos, P., Palaigeorgiou, G., Triantafyllakos, G., & Despotakis, T. (2009). Computer-based testing using "digital ink": Participatory design of a tablet PC based assessment application for secondary education. *Computers & Education*, 52, 811-819.
- Smith, C. W., & Cumming, J. (2009). *Educational assessment in the 21st century*. New York: Springer B.V.
- Sokolik, M. (2001). Computers in language teaching. In M. Celce-Murcia, (Ed.). *Teaching English as a second or foreign language*, 477-488. Boston, MA: Heinle&Heinle.
- Stevenson, J., & Gross, S. (1991). Use of a computerized adaptive testing model for ESOL/ bilingual entry/ exit decision making. In P. Dunkel(Ed.) *Computer-assisted language learning and testing*, 223-235. New York, NY: Newbury House.
- Stobart, G. (2012). Validity in formative assessment. In J. Gardner, (Ed.). *Assessment and learning*, 233-242. London: Sage Publications, Inc.
- Stricker, L. J., Wilder, G. Z., & Rock, D. A. (2004). Attitudes about the computer-based test of English as a foreign language. *Computers in Human Behavior*, 20, 37-54.
- Swanson, E. B. (1987). Information channel disposition and use. *Decision Science*, 18, 131-145.
- Taylor, C., Jamieson, J., Eignor, D., & Kirsch, I. (1998). *The relationship between computer familiarity and performance on computer-based TOEFL tasks* (TOEFL Research Report 61). Princeton, NJ: Educational Testing Service.
- Thomas, R. M. (1998). *Conducting educational research: A comparative view*. London: Bergin & Garvey.
- Tung, P. (1986). Computerized adaptive testing: Implications for language test developers. In C. W. Stansfield (Ed.). *Technology and language testing*, 11-28. Washington, DC: Teachers of English to Speakers of Other Languages.
- Vos, S. & Hacquebord, H. (1998). CAT: Computer-aided testing in the teaching of Dutch as a second language (DSL). The importance of validity. In S. Jager, J. Nerbonne, A. Essen(Eds.), *Language teaching & language technology*, 108-122. Lisse: Swets & Zeitlinger B.V.

- Vygotsky, L. S. (1978). *Mind in Society*. Cambridge, Mass: MIT Press.
- Wainer, H., & Eignor, D. (2000). Caveats, pitfalls and unexpected consequences of implementing large-scale computerized testing. In H. Wainer, N. J. Dorans, D. Eignor, R. Flaugher, B. F. Green, R. J. Mislevy, L. Steinberg, & D. Thissen (Eds.), *Computerized adaptive testing: A primer*, 271-298. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Wainer, H., & Mislevy, R. J. (2000). Item response theory, item calibration and proficiency estimation. In H. Wainer, N. J. Dorans, D. Eignor, R. Flaugher, B. F. Green, R. J. Mislevy, L. Steinberg, & D. Thissen (Eds.), *Computerized adaptive testing: A primer*, 60-100. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Walker, A. & White, G. (2013). *Technology enhanced language learning: Connecting theory and practice*. Oxford: Oxford University Press.
- Wang, S., Jiao, H., Young, M. S., Brooks, T., & Alson, J. (2007). Comparability of computer-based and paper-and-pencil testing in K-12 reading assessments: A meta-analysis of testing mode effects. *Educational and Psychological Measurement*, 68 (1), 5-24.
- Wang, H. (2010). *Comparability of computerized adaptive and paper-pencil tests*. [Online: http://images.pearsonassessments.com/images/tmrs/tmrs_rq/Bulletin_13.pdf, retrieved in August, 2013].
- Ward, W. C. (2002). Test models. In C. N. Mills, Potenza, M. T., Fremer, J. J., Ward, W. C. (Eds.), *Computer-based testing: Building the foundation for future assessments*, 37-40. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. *Language Teaching*, 31, 57-71.
- Whiston, S. C. (2009). *Principles and applications of assessment in counseling* (3rd ed.). CA: Brooks/ Cole.
- Yagcı M., Ekiz. H., ve Gelbal, S. (2011). *Çevrimiçi sınav ortamlarının öğrencilerin akademik başarılarına etkisi*. 5th international computer and instructional technologies symposium, Elazığ, Turkey, 22-24 September 2011.
- Yaman, S. O., & Cagiltay, N. E. (2010). Paper-based versus computer-based testing in engineering education. *IEEE Educon Education Engineering: The Future of Global Learning Engineering Education*, 1631-1637.
- Yunxiang, L., Ruixue, G., Lili, R., Wangjie, Quinshui, Q., & Hefei (2010). *Advantages and disadvantages of computer-based testing: A case study of service learning*. [Online: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5691870>, retrieved in July, 2013].



APPENDICES

APPENDIX 1: ETHICS COMMITTEE PERMISSION



GİZLİ
T.C.
HACETTEPE ÜNİVERSİTESİ
Rektörlük

Sayı : 35853172/ 433-2560

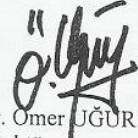
24 Ağustos 2015

EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi: 04.08.2015 tarih ve 1527 sayılı yazınız.

Enstitünüz Yabancı Diller Eğitimi Anabilim Dalı İngiliz Dili Eğitimi Bilim Dalı Tezli Yüksek Lisans Programı öğrencilerinden **Tuba ÖZTURAN**'ın Yrd. Doç. Dr. Hüseyin ÖZ danışmanlığında yürüttüğü "**Bilgisayar Temelli Ölçme-Değerlendirmenin Öğrenci Başarı ve Tutumuna Etkisi/The Impact Of Computer-Based Language Testing On Student Achievement And Attitudes In Teaching English A Foreign Language Classes**" başlıklı tez çalışması, Üniversitemiz Senatosu Etik Komisyonunun 18 Ağustos 2015 tarihinde yapmış olduğu toplantıda incelenmiş olup, etik açıdan uygun bulunmuştur.

Bilgilerinizi rica ederim.


Prof. Dr. Ömer UĞUR
Rektör a.
Rektör Yardımcısı

Ek: Tutanak

Hacettepe Üniversitesi Rektörlük 06100 Sıhhiye-Ankara
Telefon: 0 (312) 305 3001 - 3002 • Faks: 0 (312) 311 9992
E-posta: yazimd@hacettepe.edu.tr • www.hacettepe.edu.tr

Ayrıntılı Bilgi için:
Yazı İşleri Müdürlüğü
0 (312) 305 1008

APPENDIX 2: COMPUTER-ASSISTED ASSESSMENT QUESTIONNAIRE FOR THE EXPERIMENTAL GROUP

Dear Participant,

This questionnaire is administered in order to find out pre-service English teachers' opinions about computer-assisted testing. There are no right or wrong answers in this list of statements. It is simply a matter of what is true for you. Your answers will have a valuable contribution to the study. Please also make sure that the answers you give in this questionnaire will remain confidential. Please read every statement carefully and choose the best option that explains your opinion.

Thank you very much for your participation.

Section I: General Information

1. What is your gender? ☐ Female ☐ Male
2. How old are you? _____ years old.
3. What is your current Grade Point Average (GPA)? _____.
4. How many hours do you spend using the Internet in a week?
☐ Less than 3 hours ☐ 3-6 hours
☐ 6-8 hours ☐ More than 8 hours
5. How many years ago did you first begin using a computer? _____ years ago.
6. Please, indicate your overall computer literacy.
☐ 1 (very bad) ☐ 2 (bad) ☐ 3 (tolerable)
☐ 4 (good) ☐ 5 (very good)
7. What type of exam do you prefer?
☐ Computer-Assisted Exam
☐ Paper-Based Exam
☐ Both of Them (Both types should be available so that I can choose one option)

Section II: General Perceptions of Computer-Assisted Assessment

For the following items, please circle the answers that best show your opinion.

①) Strongly disagree ②) Disagree ③) Neutral ④) Agree ⑤) Strongly agree

1= Strongly Disagree 2= Disagree 3= Neutral 4= Agree 5= Strongly Agree						
	ITEMS	SD	D	N	A	SA
1.	By using the computer assisted assessment system in the exams, I was able to answer the questions more quickly compared to a paper-based exam.	①	②	③	④	⑤
2.	Using the computer assisted assessment system improved my performance in the exam.	①	②	③	④	⑤
3.	Using the computer assisted assessment system increased my productivity in the exam	①	②	③	④	⑤
4.	I found the computer assisted assessment system useful.	①	②	③	④	⑤
5.	Using computer assisted assessment system enhanced my effectiveness in the course.	①	②	③	④	⑤
6.	Computer assisted assessment system enabled me to take exams easily.	①	②	③	④	⑤
7.	I think the computer assisted assessment system was useful in Approaches to ELT course.	①	②	③	④	⑤
8.	Learning to use the computer assisted assessment system was hard for me.	①	②	③	④	⑤
9.	My interaction with the computer assisted assessment system was clear and understandable.	①	②	③	④	⑤
10.	My interaction with the computer assisted assessment system did not require a lot of mental effort.	①	②	③	④	⑤
11.	I found the computer assisted assessment system easy to use.	①	②	③	④	⑤
12.	It was easy to navigate through the computer assisted assessment system.	①	②	③	④	⑤
13.	Using the computer assisted assessment system to take exams was a good idea.	①	②	③	④	⑤
14.	Using the computer assisted assessment system to take exams was a wise idea.	①	②	③	④	⑤
15.	I liked the idea of using the computer assisted assessment system.	①	②	③	④	⑤
16.	Using the computer assisted assessment system was pleasant.	①	②	③	④	⑤
17.	The computer assisted assessment system provided an attractive exam environment.	①	②	③	④	⑤
18.	I found using the computer assisted assessment system enjoyable.	①	②	③	④	⑤
19.	In general, I was positive toward computer assisted assessment system.	①	②	③	④	⑤
20.	I intend to take courses that use the computer assisted assessment in the future.	①	②	③	④	⑤

1= Strongly Disagree 2= Disagree 3= Neutral 4= Agree 5= Strongly Agree						
	ITEMS	SD	D	N	A	SA
21.	If I am offered, I intend to take all exams with the computer assisted assessment system.	①	②	③	④	⑤
22.	I wish I used the computer assisted assessment system for other courses as well.	①	②	③	④	⑤
23.	The computer assisted assessment system was somewhat intimidating to me.	①	②	③	④	⑤
24.	I hesitated to use the computer assisted assessment system for fear of making mistakes that I couldn't correct.	①	②	③	④	⑤
25.	I felt anxious about using the computer assisted assessment system.	①	②	③	④	⑤
26.	Working with the computer assisted assessment system made me nervous.	①	②	③	④	⑤
27.	It was comfortable to work with the computer assisted assessment system.	①	②	③	④	⑤
28.	Computers are bringing us into a bright new era.	①	②	③	④	⑤
29.	The use of computers is enhancing our standard of living.	①	②	③	④	⑤
30.	There are unlimited possibilities of computer applications that have not even been thought of yet.	①	②	③	④	⑤
31.	Computers are responsible for many of the good things we enjoy.	①	②	③	④	⑤
32.	Working with computers is an enjoyable experience.	①	②	③	④	⑤
33.	I felt comfortable when using the computer assisted assessment system on my own.	①	②	③	④	⑤
34.	I was able to use the computer assisted assessment system even if there was no one around show me how to use it.	①	②	③	④	⑤
35.	I was confident that I had adequate ability to operate the computer assisted assessment system.	①	②	③	④	⑤
36.	I was confident that I could use the computer assisted assessment system even if I had no prior experience on similar systems.	①	②	③	④	⑤
37.	I could use the computer assisted assessment system if an assistant showed me how to do it first.	①	②	③	④	⑤

APPENDIX 3: COMPUTER-ASSISTED ASSESSMENT QUESTIONNAIRE FOR THE CONTROL GROUP

Dear Participant,

This questionnaire is administered in order to find out pre-service English teachers' opinions about computer-assisted testing. There are no right or wrong answers in this list of statements. It is simply a matter of what is true for you. Your answers will have a valuable contribution to the study. Please also make sure that the answers you give in this questionnaire will remain confidential. Please read every statement carefully and choose the best option that explains your opinion.

Thank you very much for your participation.

Section I: General Information

1. What is your gender? ☐ Female ☐ Male
2. How old are you? _____ years old.
3. What is your current Grade Point Average (GPA)? _____.
4. How many hours do you spend using the Internet in a week?
☐ Less than 3 hours ☐ 3-6 hours
☐ 6-8 hours ☐ More than 8 hours
5. How many years ago did you first begin using a computer? _____
years ago.
6. Please, indicate your overall computer literacy.
☐ 1 (very bad) ☐ 2 (bad) ☐ 3 (tolerable)
☐ 4 (good) ☐ 5 (very good)
7. What type of exam do you prefer?
☐ Computer-Assisted Exam
☐ Paper-Based Exam
☐ Both of Them (Both types should be available so that I can choose one option)

Section II: General Perceptions of Computer-Assisted Assessment

For the following items, please circle the answers that best show your opinion.

①) Strongly disagree ②) Disagree ③) Neutral ④) Agree ⑤) Strongly agree

1= Strongly Disagree 2= Disagree 3= Neutral 4= Agree 5= Strongly Agree						
	ITEMS	SD	D	N	A	SA
1.	By using the computer assisted assessment system in the exams, I can answer the questions more quickly compared to a paper-based exam.	①	②	③	④	⑤
2.	Using the computer assisted assessment system can improve my performance in the exam.	①	②	③	④	⑤
3.	Using the computer assisted assessment system can increase my productivity in the exam.	①	②	③	④	⑤
4.	I can find the computer assisted assessment system useful.	①	②	③	④	⑤
5.	Using the computer assisted assessment system can enhance my effectiveness in the course.	①	②	③	④	⑤
6.	Computer assisted assessment system can enable me to take exams easily.	①	②	③	④	⑤
7.	I think the computer assisted assessment system can be useful in Approaches to ELT course.	①	②	③	④	⑤
8.	Learning to use the computer assisted assessment system can be hard for me.	①	②	③	④	⑤
9.	My interaction with the computer assisted assessment system can be clear and understandable.	①	②	③	④	⑤
10.	My interaction with the computer assisted assessment system will not require a lot of mental effort.	①	②	③	④	⑤
11.	I can find the computer assisted assessment system easy to use.	①	②	③	④	⑤
12.	It will be easy to navigate through the computer assisted assessment system.	①	②	③	④	⑤
13.	Using the computer assisted assessment system to take exams is a good idea.	①	②	③	④	⑤
14.	Using the computer assisted assessment system to take exams is a wise idea.	①	②	③	④	⑤
15.	I like the idea of using the computer assisted assessment system.	①	②	③	④	⑤
16.	Using the computer assisted assessment system can be pleasant.	①	②	③	④	⑤
17.	The computer assisted assessment system can provide an attractive exam environment.	①	②	③	④	⑤
18.	I can find using the computer assisted assessment system enjoyable.	①	②	③	④	⑤
19.	In general, I am positive toward the computer assisted assessment system.	①	②	③	④	⑤
20.	I intend to take courses that use the computer assisted assessment system in the future.	①	②	③	④	⑤

1= Strongly Disagree 2= Disagree 3= Neutral 4= Agree 5= Strongly Agree						
	ITEMS	SD	D	N	A	SA
21.	If I am offered, I may take almost all exams with the computer assisted assessment system.	①	②	③	④	⑤
22.	I wish I could use the computer assisted assessment system for other courses as well.	①	②	③	④	⑤
23.	The computer assisted assessment system can be somewhat intimidating to me.	①	②	③	④	⑤
24.	I can hesitate to use the computer assisted assessment system for the fear of making mistakes that I cannot correct.	①	②	③	④	⑤
25.	I feel anxious about using the computer assisted assessment system.	①	②	③	④	⑤
26.	Working with the computer assisted assessment system can make me nervous.	①	②	③	④	⑤
27.	It can be comfortable to work with the computer assisted assessment system.	①	②	③	④	⑤
28.	Computers are bringing us into a bright new era.	①	②	③	④	⑤
29.	The use of computers is enhancing our standard of living.	①	②	③	④	⑤
30.	There are unlimited possibilities of computer applications that have not even been thought of yet.	①	②	③	④	⑤
31.	Computers are responsible for many of the good things we enjoy.	①	②	③	④	⑤
32.	Working with computers is an enjoyable experience.	①	②	③	④	⑤
33.	I can feel comfortable when using the computer assisted assessment system on my own.	①	②	③	④	⑤
34.	I can be able to use the computer assisted assessment system even if there is no one around to show me how to use it.	①	②	③	④	⑤
35.	I am confident that I have adequate ability to operate the computer assisted assessment system.	①	②	③	④	⑤
36.	I am confident that I can use the computer assisted assessment system even if I have no prior experience on similar systems.	①	②	③	④	⑤
37.	I can use the computer assisted assessment system if an assistant shows me how to do it first.	①	②	③	④	⑤

APPENDIX 4: INTERVIEW QUESTIONS

1. Have you ever taken any formal computer assisted exam before? If so, how was your reaction(s) towards it?

2. What did you like and/or dislike about the computer assisted assessment done in Approaches to ELT?

3. If you have to make a choice over a computer assisted exam and a traditional paper and pencil exam, which one would you like to take? Why?

4. What advantages do you think a computer assisted exam has over a paper-based test?

5. What disadvantages do you think a computer assisted exam has over a paper-based test?

6. Does using a computer assisted system as an assessment tool save time?

7. Do you think the computer assisted assessment (specifically web tools) should replace the traditional paper-and-pencil assessment? Why or why not?

APPENDIX 5: COURSE ACHIEVEMENT TEST

	HACETTEPE UNIVERSITY FACULTY OF EDUCATION DEPARTMENT OF FOREIGN LANGUAGE EDUCATION <i>ENGLISH LANGUAGE TEACHING PROGRAMME</i>
---	--

APPROACHES TO ENGLISH LANGUAGE TEACHING **MODULE 1 – ACHIEVEMENT TEST**

NAME AND SURNAME	:
STUDENT ID NUMBER	:
SECTION	:
SIGNATURE	:
SCORE	:

1. Communicative Language Teaching aims to apply the theoretical perspective of the communicative competence by _____.

- A) acknowledging the independence of language and communication
- B) enabling learners to know the rules of language
- C) making communicative competence the goal of language teaching
- D) enhancing the memorization of fixed routines

2. The modal verb “MAY” is contextually used to “make a prediction” or to “give permission,” which in this case means _____.

- A) a single form of language can be used for a variety set of functions
- B) a variety of forms can be used to accomplish a single function
- C) a variety of forms can be used for describing different sets of functions
- D) the sequence of units is determined solely by principles of linguistic complexity

3. One of the obvious characteristics of Communicative Language Teaching implies that almost everything done in the classroom is done with _____.

- A) chain drill
- B) authentic material
- C) communicative intent
- D) structured feedback

4. In Communicative Language Teaching, language learners are expected to _____.

- A) get the meaning in the target language through actions
- B) to memorize some part of drills or dialogues
- C) master in learning forms of language
- D) communicate using the language

5. Taking into consideration Communicative Language Teaching principles, language learners use the target language to specifically accomplish linguistic _____.

- A) functions
- B) collocations
- C) patterns
- D) clarifications

If two speakers know that it is 10 o'clock and one asks the other, “What time is it?” and the other speaker answers, “10 O'clock.” This is not really communicative.

6. What feature of a communicative exchange or activity does the above statement illustrate?

- A) Choice
- B) Feedback
- C) Information Gap
- D) Grammaticality

7. Communicative Language Teaching does NOT demand that ESL/EFL learners use the target language a great deal via implementing a technique called _____.

- A) role plays
- B) transformation drills
- C) language games
- D) problem-solving tasks

8. According to Communicative Language Teaching, the target language serves as _____.

- I an object to be studied
- II a vehicle for classroom communication
- III a model to facilitate the native language acquisition
- IV a tool to use authentic materials
- A) Only III
- B) I - II - III - IV
- C) I - II - III
- D) I - II - IV

9. The one important feature of Communicative Language Teaching is that it specifically attends to _____ in second/foreign language learning contexts.

- A) negotiating meaning
- B) developing inner criteria
- C) overcoming native language habits
- D) acquiring native-speaker-like pronunciation

10. In Communicative Language Teaching, it is also important for language students to efficiently learn about _____ and _____; that is, those properties of language which bind the sentences together.

- A) form / content
- B) cohesion / coherence
- C) notion / function
- D) competence / performance

11. According to Communicative Language Teaching, language learners should specifically work on _____ level.

- A) phrasal or sentential
- B) sentential or discourse
- C) discourse or suprasentential
- D) suprasentential or phrasal

12. In Communicative Language Teaching, language teachers should _____ errors and _____ them as a natural outcome of communicative skills development.

- A) reject / prevent
- B) prevent / consider
- C) prohibit / value
- D) put up with / view

13. Language teachers who support the principles of Communicative Language Teaching primarily perform the role of a(n) _____ during their classroom activities.

- A) director and authority
- B) observer and councilor
- C) counselor and director
- D) advisor and facilitator

14. In the language classrooms of teachers who conform to the principles of Communicative Language Teaching, it is often possible to notice the students specifically exercising a choice about _____ and _____ to say things.

- A) what / how
- B) when / where
- C) how / when
- D) where / what

15. The relevant grammar and vocabulary the students learn from language contexts do not essentially follow from _____.

- A) language functions
- B) sequence of actions
- C) situational contexts
- D) roles of interlocutors

16. Communicatively competent language learners particularly include those who can linguistically use _____.

- I. language appropriate to a social context
- II. many different language forms to perform a function
- III. a single form to perform a variety of functions
- IV. native-like fluency in daily communication

- A) I - II
- B) I - III
- C) I - II - III
- D) I - II - III - IV

17. The truly communicative activities implemented in language classrooms are NOT characterized by the specific feature of _____.

- A) choice
- B) well-formedness
- C) feedback
- D) information gap

18. Which of the following statement(s) can be defined as part(s) of communicative competence?

- I. to know the variation of language according to the setting
- II. to use a language without error(s)
- III. to use a language in a different range of purposes

- A) I – III
- B) II - III
- C) I - II
- D) I – II – III

19. Communicative Language Teaching does not necessarily stress _____ over _____ in ESL/EFL language classrooms.

- A) discourse level / sentence level
- B) verbal behavior / nonverbal behavior
- C) language functions / language forms
- D) communicative competence / linguistic competence

20. Students undergoing language instruction by Communicative Language Teaching principles work on the skill(s) of _____ in the classroom from the very beginning.

- A) speaking
- B) speaking and listening
- C) speaking, listening, and reading
- D) speaking, listening, reading, and writing

21. Communicative Language Teaching typically presents the content in language classrooms specifically through a _____ type of syllabus.

- A) lexical
- B) structural
- C) combined
- D) functional

22. Language teachers are asked to use the target language _____ in CLT-based classrooms.

- I. transformation drills
- II. on assigning homework
- III. as an object to be studied
- IV. during communicative activities

- A) I – II
- B) II – III – IV
- C) II - III
- D) I – III - IV

23. In Communicative Language Teaching, language teachers encourage the use of authentic language materials to specifically help language students _____.

- I. realize the cohesion and coherence properties of language
- II. transfer what they learn in the classroom to the outside world
- III. get categorically exposed to natural language in a variety of situations
- IV. get cultural information about the target language

- A) II-III-IV
- B) I-II
- C) I-II-III
- D) I-II-III-IV

24. In Communicative Language Teaching it is believed that the unscrambling scrambled sentences directly teaches language learners about the _____ and _____ properties of language.

- A) use – usage
- B) cohesion – coherence
- C) forms - functions
- D) vocabulary – structures

25. Language learners' communicative competence consists of the knowledge of _____.

- I. forms
- II. meanings
- III. functions

- A) Only II
- B) Only III
- C) II – III
- D) I – II – III

26. The language user's usage of the target language appropriately based on the situation is accounted for to inherently make up the _____ competence.

- A) communicative
- B) linguistic
- C) discourse
- D) strategic

27. In Communicative Language Teaching, meaning exclusively _____.

- A) resides in the text documented by the writer for the reader
- B) comes off through translation done by the teacher
- C) arises through negotiation between the reader and the writer
- D) depends on practicing a grammar point in context, using lots of repetition

28. Communicative Language Teaching teachers are often seen establishing situations to _____ communication between and among the language learners.

- A) assess
- B) prompt
- C) modify
- D) substitute

29. In Communicative Language Teaching, language activities in classrooms are characteristically initiated from _____ to that of _____.

- I. groups - teachers
- II. students – teachers
- III. teachers – students
- IV. students – students

- A) I-II
- B) II-III
- C) II-III-IV
- D) I-III-IV

30. Teachers of Communicative Language Teaching should avoid assuming the role of a(n) _____ in second/foreign language classrooms.

- A) advisor answering students' questions
- B) monitor of students' performance
- C) facilitator of their students' learning
- D) director of all students' behaviors

31. Students' success in language classrooms conducted by Communicative Language Teaching principles is determined by the crucial parameters of _____, acquired through language learning.

- A) fluency and structure
- B) accuracy and vocabulary
- C) fluency and accuracy
- D) structure and vocabulary

32. In Communicative Language Teaching classrooms, an outsider may see the teacher distributing a handout containing a sports column elicited from a recent newspaper to intentionally introduce _____ language in a real context.

- A) authentic
- B) reported
- C) semantic
- D) written

33. ESL/EFL students in CLT-directed classrooms interact a great deal with one another in a configuration of _____.

- I. Individual work
- II. pairs or triads
- III. whole group

- A) Only I
- B) I - II
- C) II – III
- D) I – II – III

34. _____ syllabus cannot be regarded as a communicative syllabus design.

- A) An ESP (English for Specific Purposes)
- B) A grammatical syllabus
- C) A functional syllabus
- D) A skill-based syllabus

35. Exercises in many Communicative Language Teaching course books take students from _____ to _____ to communicative practice.

- A) mechanical / meaningful
- B) meaningful / structural
- C) notional / meaningful
- D) structural / notional

36. Communicative Language Teaching introduces new activity types for language classes; one of which is the concept of _____ that refers to the usage of the target language for completing an objective. Puzzles, games, map reading are some examples of it.

- A) jig-saw
- B) information-gap
- C) opinion-sharing
- D) task-completion

37. It would be incorrect to claim that classroom interaction in communicative language teaching is heavily characterized by _____.

- A) encouraging the development of cooperative relationships
- B) enhancing the memorization of fixed routines
- C) introducing authentic texts into the learning situation
- D) granting students an opportunity to work on negotiating meaning

38. Learning to use language forms appropriately is considered to serve as a key feature of _____ competence.

- A) communicative
- B) lexical
- C) strategic
- D) linguistic

39. The selection and sequence of language within the content-based scheme items is specifically made in terms of _____.

- A) strategic competence
- B) pre-determined syllabi
- C) communicative needs
- D) structure dependency

40. A language teacher should make use of _____ in class to _____ success in CLT-based language classes.

- A) action sequence/ diminish
- B) pair work/ decrease
- C) language chunks/ boost
- D) collaborative learning/ trigger

41. _____ is an important aid for language teachers in CLT-based classes to help language learners communicate meaningfully.

- A) Role reversal
- B) Jig-saw activity
- C) Peripheral learning
- D) Novelty

42. _____ can be a part of materials used in Communicative Language Teaching classes as they reflect the _____ language.

- A) Newspapers/ authentic
- B) Magazines/ native
- C) Drills/ target
- D) Charts/ second

43. _____ is a newly emerged type of assessment with the CLT movement so as to build up a/an _____ picture of what language learners can do in the target language.

- A) Gap-filling/ detailed
- B) Multiple-choice/ overall
- C) Observation/ comprehensive
- D) Peer correction/ elaborated

44. In Communicative Language Teaching, language learners will learn to communicate _____.

- A) by practicing notional and socially appropriate language.
- B) by practicing notional and linguistically appropriate language
- C) by practicing functional and socially appropriate language
- D) by practicing structural and linguistically appropriate language

45. According to Communicative Language Teaching, knowledge of the forms of language _____.

- A) Is sufficient for an effective communication process.
- B) can by no means be regarded as the core of CLT
- C) is supposed to be a base for communicative competence
- D) is insufficient to enhance grammatical competence

46. Usage of literary texts in the target language in CLT-based language classes can greatly help language learners because _____.

- A) they commonly introduce cultural properties
- B) they make the learners familiar with specific language structures
- C) they are easy to be read and worked out by learners
- D) they offer role-reversal opportunities

47. Using the principles of Communicative Language Teaching effectively in language classes, language teachers should use _____ and _____ tasks commonly.

- A) fluency-based / notion-based
- B) functional / structural
- C) notion-based / functional
- D) fluency/ accuracy

48. _____ is commonly _____ to enhance a motivating classroom atmosphere for language learners in CLT-directed classrooms.
- A) Individual training / assumed
 - B) Cooperative learning / utilized
 - C) Peer correction / disposed
 - D) Group work/ refrained
49. According to the principles of Communicative Language Teaching, people can learn a language best if they are taught through truly _____ activities.
- A) meaningful
 - B) comprehensible
 - C) informative
 - D) exceptional
50. The target language is supposed to be _____ other subjects in Communicative Language Teaching classes in order to make language learners explore the _____ outside the class.
- A) combined with / linguistic notions
 - B) separated from / classroom language
 - C) integrated with / real language
 - D) incorporated into / contrasting functions
51. The _____ version of Communicative Language Teaching claims that language is acquired through communication, so that it is not merely a question of activating an existing but inert knowledge of language, but of simulating the development of the language system itself.
52. The _____ language based on Communicative Language Teaching is viewed as serving as that of a vehicle for classroom communication.
53. Using an excerpt and letting the language learners discover the rules mean _____ learning.
54. In Communicative Language Teaching classes, communication is a/an _____ process that often calls upon the use of several language skills or modalities.
55. _____ practice refers to activities where practice in using language within a real communicative context is the focus, where real information is exchanged, and where the language used is not totally predictable.
56. _____ practice refers to a controlled practice activity which students can successfully carry out without necessarily understanding the language they are using.
57. _____ practice refers to an activity where language control is still provided but where students are required to make meaningful choices when carrying out practice.
58. One of the most characteristic features of CLT is that it pays systematic attention to functional as well as _____ aspects of language.
59. The _____ version of Communicative Language Teaching stresses the importance of providing learners with opportunities to use their English for communicative purposes and attempts to integrate such activities into a wider programme of language teaching.

60. Communicative Language Teaching maintains that one linguistic form can accordingly have many different



APPENDIX 6: RELIABILITY COEFFICIENTS OF THE ITEMS IN COMPUTER- ASSISTED ASSESSMENT QUESTIONNAIRE FOR THE EXPERIMENTAL GROUP

<i>Item</i>	<i>Item Description</i>	<i>r</i>
	FACTOR I INTENTION (INT)	.963
1.	By using the computer assisted assessment system in the exams, I was able to answer the questions more quickly compared to a paper-based exam.	.963
6.	Computer assistedassessment system enabled me to take exams easily.	.960
13.	Using the computer assisted assessment system to take exams was a good idea.	.958
14.	Using the computer assistedassessment system to take exams was a wise idea.	.961
15.	I liked the idea of using the computer assistedassessment system.	.957
16.	Using the computer assistedassessment system was pleasant.	.959
17.	The computer assistedassessment system provided an attractive exam environment.	.959
18.	I found using the computer assistedassessment system enjoyable.	.963
19.	In general, I was positive toward computer assistedassessment system.	.960
20.	I intend to take courses that use the computer assistedassessment in the future.	.957
27.	It was comfortable to work with the computer assistedassessment system.	.961
	FACTOR II PERCEIVED USEFULNESS (PU)	.920
2.	Using computer assistedassessment system improved my performance in the exam.	.895
3.	Using computer assistedassessment system increased my productivity in the exam	.911
5	Using computer assistedassessment system enhanced my effectiveness in the course.	.901
21.	If I am offered, I intend to take all exams with the computer assistedassessment system.	.902
22.	I wish I used the computer assistedassessment system for other courses as well.	.898
	FACTOR III PERCEIVED EASE OF USE (PEOU)	.812
4.	I found the computer assistedassessment system useful.	.785
7.	I think the computer assistedassessment system was useful in <i>Approaches to ELT</i> course.	.775
8.	Learning to use the computer assistedassessment system was hard for me.	.842
9.	My interaction with the computer assistedsystem was clear and understandable.	.787
10.	My interaction with the computer assistedassessment system did not require a lot of mental effort.	.820
11.	I found the computer assistedassessment system easy to use.	.786
12.	It was easy to navigate through the computer assistedassessment system.	.792
33.	I felt comfortable when using the computer assistedassessment system on my own.	.783
34.	I was able to use the computer assistedassessment system even if there was no one around show me how to use it.	.793
35.	I was confident that I had adequate ability to operate the computer assistedassessment system.	.777
36.	I was confident that I could use the computer assistedassessment system even if I had no prior experience on similar systems.	.791
37.	I could use the computer assistedassessment system if an assistant showed me how to do it first.	.836
	FACTOR IV ANXIETY (ANX)	.771
23.	The computer assistedassessment system was somewhat intimidating to me.	.778

24.	I hesitated to use the computer assistedassessment system for fear of making mistakes that I couldn't correct.	.723
25.	I felt anxious about using the computer assistedassessment system.	.622
26.	Working with the computer assistedassessment system made me nervous.	.720
	FACTOR V COMPUTER ATTITUDE (CA)	.859
28.	Computers are bringing us into a bright new era.	.822
29.	The use of computers is enhancing our standard of living.	.850
30.	There are unlimited possibilities of computer applications that have not even been thought of yet.	.863
31.	Computers are responsible for many of the good things we enjoy.	.821
32.	Working with computers is an enjoyable experience.	.780



APPENDIX 7: RELIABILITY COEFFICIENTS OF THE ITEMS IN COMPUTER- ASSISTED ASSESSMENT QUESTIONNAIRE FOR THE CONTROL GROUP

<i>Item</i>	<i>Item Description</i>	<i>r</i>
	FACTOR I INTENTION (INT)	.939
1.	By using the computer assistedassessment system in the exams, I can answer the questions more quickly compared to a paper-based exam.	.937
6.	Computer assistedassessment system can enable me to take exams easily.	.938
13.	Using the computer assistedassessment system to take exams is a good idea.	.926
14.	Using the computer assistedassessment system to take exams is a wise idea.	.931
15.	I like the idea of using the computer assistedassessment system.	.930
16.	Using the computer assistedassessment system is pleasant.	.931
17.	The computer assistedassessment system provides an attractive exam environment.	.941
18.	I can find using the computer assistedassessment system enjoyable.	.932
19.	In general, I am positive toward computer assistedassessment system.	.934
20.	I intend to take courses that use the computer assistedassessment in the future.	.937
27.	It is comfortable to work with the computer assistedassessment system.	.933
	FACTOR II PERCEIVED USEFULNESS (PU)	.842
2.	Using the computer assistedassessment system can improve my performance in the exam.	.785
3.	Using the computer assistedassessment system can increase my productivity in the exam.	.810
5.	Using the computer assistedassessment system can enhance my effectiveness in the course.	.828
21.	If I am offered, I intend to take all exams with the computer assistedassessment system.	.802
22.	I wish I could use the computer assistedassessment system for other courses as well.	.827
	FACTOR III PERCEIVED EASE OF USE (PEOU)	.875
4.	I can find the computer assistedassessment system useful.	.860
7.	I think the computer assistedassessment system can be useful in Approaches to ELT course.	.873
8.	Learning to use the computer assistedassessment system can be hard for me.	.861
9.	My interaction with the computer assistedassessment system can be clear and understandable.	.861
10.	My interaction with the computer assistedassessment system will not require a lot of mental effort.	.881
11.	I can find the computer assistedassessment system easy to use.	.857
12.	It will be easy to navigate through the computer assistedassessment system.	.859
33.	I can feel comfortable when using the computer assistedassessment system on my own.	.867
34.	I can use the computer assistedassessment system even if there is no one around show me how to use it.	.858
35.	I am confident that I have adequate ability to operate the computer assistedassessment system.	.853
36.	I am confident that I can use the computer assistedassessment system even if I have no prior experience on similar systems.	.859
37.	I can use the computer assistedassessment system if an assistant shows me how to do it first.	.881
	FACTOR IV ANXIETY (ANX)	.596

23.	The computer assistedassessment system is somewhat intimidating to me.	.744
24.	I hesitate to use the computer assistedassessment system for fear of making mistakes that I can't correct.	.447
25.	I feel anxious about using the computer assistedassessment system.	.418
26.	Working with the computer assistedassessment system makes me nervous.	.393
	FACTOR V COMPUTER ATTITUDE (CA)	.824
28.	Computers are bringing us into a bright new era.	.799
29.	The use of computers is enhancing our standard of living.	.748
30.	There are unlimited possibilities of computer applications that have not even been thought of yet.	.799
31.	Computers are responsible for many of the good things we enjoy.	.805
32.	Working with computers is an enjoyable experience.	.790



APPENDIX 8: ORIGINALITY REPORT

Tuba ÖZTURAN'ın "THE IMPACT OF COMPUTER ASSISTED ASSESSMENT ON THE EXAM SUCESS AND ATTITUDES OF PROSPECTIVE ENGLISH TEACHERS" başlıklı yüksek lisans tez çalışmasına ait orijinallik raporu TurnItIn uygulamasında hazırlanmış ve %14 olarak bulunmuştur. Bu orijinallik yüzdesi kaynakça bölümü hariç ve alıntılar dâhil olarak hesaplanmıştır.

Turnitin Document Viewer - Google Chrome
https://www.turnitin.com/dv?u=1025212600&lang=en_us&o=628193758&s=3

MA Theses Thesis Submission-Tuba Özturan - DUE.z...

Originality GradeMark PeerMark

Thesis Submission-Tuba Özturan
BY TUBA ÖZTURAN

turnitin 14% SIMILAR OUT OF 0

Match Overview

1	Oz, Hüseyin. "Pre-serv...	2%
2	Submitted to Middle E...	2%
3	Submitted to Techno...	1%
4	CHAU, PATRICK Y.K. ...	1%
5	ijjalw.org	<1%
6	etd.lib.ttu.edu	<1%
7	Submitted to Bilkent U...	<1%
8	Submitted to Macquari...	<1%
9	www.tojet.net	<1%

THE IMPACT OF COMPUTER ASSISTED ASSESSMENT
ON THE EXAM SUCESS AND ATTITUDES OF
PROSPECTIVE ENGLISH TEACHERS

BİLGİSAYAR TEMELLİ ÖLÇME-DEĞERLENDİRMEİNİN
İNGİLİZCE ÖĞRETMEN ADAYLARININ BAŞARI VE
TUTUMU ÜZERİNE ETKİSİ

Tuba ÖZTURAN

AUTHOR CV

Kişisel Bilgiler

Adı Soyadı	Tuba Özturan
Doğum Yeri	Erzincan
Doğum Tarihi	18.10.1988

Eğitim Durumu

Lise	Milliyet Anadolu Öğretmen Lisesi Erzincan/ Türkiye	2006
Lisans	Gazi Üniversitesi İngilizce Öğretmenliği Ankara/ Türkiye	2010
Yüksek Lisans		
Yabancı Dil	İngilizce: Okuma (Çokiyi), Yazma (Çokiyi), Konuşma (Çokiyi), Dinleme (Çokiyi)	

İş Deneyimi

Stajlar	Rıdvan Ege İÖO Ankara/ Türkiye	Ekim, 2009- Mayıs, 2010
Projeler	-	-
Çalıştığı Kurumlar	Erzincan Üniversitesi	Ocak 2011-

Akademik Çalışmalar

Yayınlar (Ulusal, uluslararası makale, bildiri, poster vb gibi.)

Sarıçoban, A. & Özturan, T. (2012). Vocabulary Learning on Move: An Investigation of Mobile Assisted Vocabulary Learning Effect over Students' Success and Attitudes, <i>6th International Computer and Instructional Technologies Symposium</i> , Gaziantep, Turkey
Öz, H. & Özturan, T. (2013). Comparability, Reliability and Validity of Paper and Computer-based Achievement Tests: A Study of Test Mode Effect, <i>1st International ELT Conference</i> , Ankara, Turkey
Sarıçoban, A. & Özturan, T. (2013). Vocabulary Learning on Move: An Investigation of Mobile Assisted Vocabulary Learning Effect over Students' Success and Attitudes, <i>EKEV Akademi, Dergisi</i> , 17 (54), 213-224

Seminer ve Çalıştaylar

-

Sertifikalar

-

İletişim

e-Posta Adresi	tbozturan@gmail.com
Jüri Tarihi	28.01.2016